

GenCore version 5.1.3

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 19:04:17 ; Search time 4194.54 Seconds
(without alignments)
10726.543 Million cell updates/sec

Title: US-09-888-358-2

Perfect score: 1546

Sequence: 1 GGCTAGGTGCGCTGCGAGCG.....GCTAAAGATGATGAATTC 1546

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

To number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenEmbl:**

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.em.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1546	100.0	1546	6	AX403085
2	1546	100.0	1546	9	AF151827
3	1504.8	97.3	1579	9	BC001398
4	1498	96.9	1560	9	BC009330
5	1497	96.8	1550	9	AL133584
6	1493.8	96.6	1554	9	AK026060
7	1443.8	93.4	1643	6	AR116038
8	1381.2	89.3	1402	17	AF119864
9	1054.4	68.2	1114	6	AX403084
10	1054.4	68.2	1114	9	AF317711
11	786.8	50.9	173434	2	AC098934
12	551.2	35.7	195701	2	AC124203
13	537.2	34.7	156175	2	AC111356
14	530	34.3	128802	2	AC123139
15	522	33.8	537	6	AX106606
16	522	33.8	537	6	AX140897
17	522	33.8	537	6	AX200757
18	522	33.8	537	6	AX267413
19	520	33.6	55252	2	AC119489
20	494	32.0	139488	9	AC003043
21	494	32.0	180520	2	AC019152
22	470.4	30.4	189391	2	AC019151
23	458.4	29.7	177613	2	AC123014
24	445.4	28.8	452	6	AX408486
25	437.4	28.3	758	6	AX421524
26	406.6	26.3	194025	2	AC114721
27	369	23.9	384	9	BC011178
28	296.8	19.2	138839	2	AC108295
29	242.2	15.7	1284	9	BC027322
30	240.6	15.6	1517	9	AF125531
31	218.4	14.1	240	6	AX408532
32	210.4	13.6	162456	2	AC099097
33	174.8	11.3	55774	2	AC091303
34	170.2	11.0	512	6	AX400653
35	166.6	10.8	82586	2	AC096040
36	166.6	10.8	88670	2	AC127606
37	154.4	10.0	194025	2	AC114721
38	143.8	9.3	195768	2	AC068775
39	140.4	9.1	165490	2	AC025326
40	140.4	9.1	190430	2	AC087873
41	140.4	9.1	219037	10	AL596258
42	139.6	9.0	197197	2	AC118439
43	138.8	9.0	217422	2	AC116713
44	130.4	8.4	3070	3	AY070516
45	130.4	8.4	3070	3	AY070516

ALIGNMENTS

RESULT 1
AX403085

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AX403085 Sequence 2 from Patent WO0198355. 1546 bp DNA linear PAT 07-JUN-2002

AX403085

AX403085.1 GI:21388034

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Lewin,D., Adams,S.H. and Yu,X.X.

Cgi-69 compositions and methods of use

Patent: WO 0198355-A 27-DEC-2001;


```

/codon_start=1
/product="CGI-69 protein"
/protein_id="AA034064.1"
/db_xref="GI:4929607"
/translation="MADQDPAGISPLQOMVASGTGAVVTSLEWPLDIVKVELQSREP
SVASLSPSSRLSLSYTKKLLYCNQVLYPLDNGMRCATWFDQPFRTGMDA
FKVIRHSGKRLWSGPAFLVNTVPAIYFYAYDOLKAFILGRLALISDLYAPMAG
ALUARIGTVTVISPLRLKQLQAVSYRELQACVTAQVAGWSRLWLGWPTALRD
VPSALYMFNVELVLEKMLNGFRPKDQTSVGNFSVAGGISGTVAVALTLTLPDVKTORQ
VALGMAEAVRNPLHVDSTWLLIRIRAESCTGKLPAGLPRLIIKAAPSCAIMISTYVE
FKGSFFQSLNDRLLGG"
BASE COUNT      282 a  489 c  459 g  326 t
ORIGIN

```

```

Query Match      100.0%; Score 1546; DB 9; Length 1556;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTAGGTGCGCTCGAGCGCGGAGCCAGCAGCGGCGGAGCGAGTAAATGGCGCGCT 60
DB 1 GGCTAGGTGCGCTCGAGCGCGGAGCCAGCGGCGGAGCGGAGTAAATGGCGCGCT 60

QY 61 GGCCTGGGCGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120
DB 61 GGCCTGGGCGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120

QY 121 GCTGACGAGGACCTCGCGGAGCAGCGCCCTCCAGCAATGGTGGCTCAGCAGCGG 180
DB 121 GCTGACGAGGACCTCGCGGAGCAGCGCCCTCCAGCAATGGTGGCTCAGCAGCGG 180

QY 181 GCTGTGTTACCTCTCTCTATGACACCCCTCGAGCGTGGTGAAGGTTTCGCGTCT 240
DB 181 GCTGTGTTACCTCTCTCTATGACACCCCTCGAGCGTGGTGAAGGTTTCGCGTCT 240

QY 241 CAGCGCGCTTCATGCGCAGCGAGCTGATGCTTCCTCCAGACTGTGAGCGCTTCCTAT 300
DB 241 CAGCGCGCTTCATGCGCAGCGAGCTGATGCTTCCTCCAGACTGTGAGCGCTTCCTAT 300

QY 301 ACCAAATGGAAGTGCCCTCTGTAATGGTGTCTTGAGCGCTCTGTACCTGTGCCCA 360
DB 301 ACCAAATGGAAGTGCCCTCTGTAATGGTGTCTTGAGCGCTCTGTACCTGTGCCCA 360

QY 361 AATGGTGGCGCTGTGCGACCTGTTTCAAGACCTTACCGCTTCACTGGCAGCATGGAT 420
DB 361 AATGGTGGCGCTGTGCGACCTGTTTCAAGACCTTACCGCTTCACTGGCAGCATGGAT 420

QY 421 GCCTTCGTGAAGATCGTAGGCAAGGCGACAGGACCCCTCTGAGCGGCTCCCGGC 480
DB 421 GCCTTCGTGAAGATCGTAGGCAAGGCGACAGGACCCCTCTGAGCGGCTCCCGGC 480

QY 481 ACCCTGTGATGATGTCGCGAGCTACCGCATCTACTTCAGTCCATATGACCAACTGAAG 540
DB 481 ACCCTGTGATGATGTCGCGAGCTACCGCATCTACTTCAGTCCATATGACCAACTGAAG 540

QY 541 GCCTTCCTGTGTGTCGAGCGCTGACCTCTGACCTCTAGCAGCACCCTGTTGGTGGCGG 600
DB 541 GCCTTCCTGTGTGTCGAGCGCTGACCTCTGACCTCTAGCAGCACCCTGTTGGTGGCGG 600

QY 601 CTGCGCGCGCTGCGGCAACCGTGTGATCAGCCCCCTGGAGCTTATCGGCAAAAGCTG 660
DB 601 CTGCGCGCGCTGCGGCAACCGTGTGATCAGCCCCCTGGAGCTTATCGGCAAAAGCTG 660

QY 661 CAGGCTCAGCATGTGTCGAGCGAGCTGGTGGCTGTCGAGCTGTCGAGCTGTCGAG 720
DB 661 CAGGCTCAGCATGTGTCGAGCGAGCTGGTGGCTGTCGAGCTGTCGAGCTGTCGAG 720

QY 721 GGTGCTGGCGCTCACTGTGCTGGCTGGCGGCCCACTGCGCTTCGAGATGTGCGCTTC 780
DB 721 GGTGCTGGCGCTCACTGTGCTGGCTGGCGGCCCACTGCGCTTCGAGATGTGCGCTTC 780

QY 781 TCAGCGCTGACTGTTCACTATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTCAGCGCG 840
DB 781 TCAGCGCTGACTGTTCACTATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTCAGCGCG 840

```

```

QY 841 AAGGACCAAGACTTCTGTGGGATGAGCTTTGTGGCTGTGGCTATCTCAGGACCGTGGCT 900
DB 841 AAGGACCAAGACTTCTGTGGGATGAGCTTTGTGGCTGTGGCTATCTCAGGACCGTGGCT 900

QY 901 GAGTGTGCTGACTCTACCTTTTGTGGTGAAGACCCCAAGCCAGGTGCTCTGTGGAGCG 960
DB 901 GAGTGTGCTGACTCTACCTTTTGTGGTGAAGACCCCAAGCCAGGTGCTCTGTGGAGCG 960

QY 961 ATGGAGGCTGTGAGAGTGAAACCCCTTCATGTGACTTCCACCTGGCTGGCTCTGCGGAGG 1020
DB 961 ATGGAGGCTGTGAGAGTGAAACCCCTTCATGTGACTTCCACCTGGCTGGCTCTGCGGAGG 1020

QY 1021 ATCCGGGCGGAGTGGGACCAAGGAGACTTTTCAGGCTTCTTCTCGGATCATCAAG 1080
DB 1021 ATCCGGGCGGAGTGGGACCAAGGAGACTTTTCAGGCTTCTTCTCGGATCATCAAG 1080

QY 1081 GCTGCCCCCTCTCTGTCATCATGATCAGACCTATCAGTTCGGGAAAAGCTTCTCCAG 1140
DB 1081 GCTGCCCCCTCTCTGTCATCATGATCAGACCTATGATTTGGCAAAAGCTTCTCCAG 1140

QY 1141 AGGCTGAACCAAGACCGGCTTCTGGCGGCTGAAGGGGCAAGGAGCAAGACCCGTC 1200
DB 1141 AGGCTGAACCAAGACCGGCTTCTGGCGGCTGAAGGGGCAAGGAGCAAGACCCGTC 1200

QY 1201 TCTCCACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 TCTCCACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260

QY 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320

QY 1321 CTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
DB 1321 CTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380

QY 1381 ATCACTTACCAACCCCAAGTTCAGACCAAAATCTTCAAGCTGCCCCCTTCTGTTTCC 1440
DB 1381 ATCACTTACCAACCCCAAGTTCAGACCAAAATCTTCAAGCTGCCCCCTTCTGTTTCC 1440

QY 1441 CTGTGTTGTGTAGTGTGGGATGTCTCCAGGAAACCAAGAGCCCTCAGCCCTGTGTAGT 1500
DB 1441 CTGTGTTGTGTAGTGTGGGATGTCTCCAGGAAACCAAGAGCCCTCAGCCCTGTGTAGT 1500

QY 1501 CTCCTGACCTTCTTAATTTCTTAAGTCTAAAGATGATGAATTC 1546
DB 1501 CTCCTGACCTTCTTAATTTCTTAAGTCTAAAGATGATGAATTC 1546

```

```

RESULT 3
BC001398
LOCUS      1579 bp      mRNA      linear      PRI 12-JUL-2001
DEFINITION Homo sapiens, CGI-69 protein, clone MGC:1163 IMAGE:3051436, mRNA,
complete cds.
ACCESSION BC001398
VERSION    BC001398.1 GI:12655092
KEYWORDS   MGC.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1579)
AUTHORS    Strausberg,R.
TITLE      Direct Submission
JOURNAL
SUBMITTED (12-DEC-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT    Contact: MGC help desk
           Email: cgabs@mail.nih.gov
           Tissue Procurement: ATCC

```

cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mc@hgrl.nih.gov
 Shevchenko, Y., Weherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.B., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, E., Stantripop, S., Thomas, P.J.,
 T'ongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 4 Row: 0 Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10438782.

FEATURE:

Source

Location/Qualifiers

1..1579

/organism="Homo sapiens"

/db_xref="LocusID:51629"

/db_xref="taxon:9606"

/clone="MGC:1163 IMAGE:3051436"

/tissue type="Colon, adenocarcinoma"

/clone_lib="NIH MGC 15"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

131..1186

/codon_start=1

/product="CGI-69 protein"

/protein_id="AAH01398.1"

/db_xref="GI:12655093"

/translation="MADQDPAGISPLQOMVAGTGAIVTSLMTPLDVVKVLOSQR
 SMASLPPSLRLWSLYTKWKLCLYCNGLVLEPLCNGARCATWPDPTFTGMDA
 FVKIVREGRTLWSGLPATLVMYTPATAYFTAYDLKAFGLGRALTSIDLYAMVAG
 ALARLGTIVISPLEMRKGAQHVSTRELGCRTAVAGQWRLGSLWGLGFWPTALRD
 VPFSAIWFNIVELVSLWGLFDPKQTSVGVSVFVAGISGTVAALVLPFDWVKTORQ
 VALGAEAVRNPLPHVDSTLWLLRRIRAESGTGLFAGFLPRIIAKAPSCALMISTVE
 FGKSFQRLNQDRLGG"

BASE COUNT 293 a 496 c 464 g 326 t

ORIGIN

Query Match 97.3%; Score 1504.8; DB 9; Length 1579;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 17; Indels 4; Gaps 1;
 1 GGCTAGGTGCGTGGAGCGCGG---GAGCCACGAGCGGCGGACGAGCTAATGGGCC 56
 Db 10 GCTAGGTGCGTGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 69
 QY 57 GCTGGCCCTGGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 116
 Db 70 GCTGACCTGGGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 129
 QY 117 GATGCTGACAGGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 176
 Db 130 GATGCTGACAGGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 189
 QY 177 CGGGCGCTGTGTTTACCTCTCTCTCATGACCCCTGAGGTGAGGTGAGGTGAGGTG 236
 Db 190 CGGGCGCTGTGTTTACCTCTCTCTCATGACCCCTGAGGTGAGGTGAGGTGAGGT 249
 QY 237 GTCTCAGCGCGCCCTTCCATGGCAGCGAGCTGATGCTTCTCTCCAGACTGTGGAGCCTCTC 296
 Db 250 GTCTCAGCGCGCCCTTCCATGGCAGCGAGCTGATGCTTCTCTCCAGACTGTGGAGCCTCTC 309
 QY 297 CTATACCAATGAGAGTGCCTCTCTGATTTGCAATGGTGTCTGGAGCCTCTGATCTGTG 356
 Db 310 CTATACCAATGAGAGTGCCTCTCTGATTTGCAATGGTGTCTGGAGCCTCTGATCTGTG 369

QY 357 CCAAAATGGTCCCGCTGTCACCTGGTTCACAGACCTACCGGCTTCACTGGCACCAT 416
 Db 370 CCAAAATGGTCCCGCTGTCACCTGGTTCACAGACCTACCGGCTTCACTGGCACCAT 429
 QY 417 GGATGCTTCTGTGAAGATCGTAGGACAGGAGGACAGGAGGACCTCTGGAGGCGGCTCCC 476
 Db 430 GGATGCTTCTGTGAAGATCGTAGGACAGGAGGACAGGAGGACCTCTGGAGGCGGCTCCC 489
 QY 477 CGCAGCCTGTGTGATGATGTGTCAGCTACCGGACCTTACTTTCAGTGCCTATGACCAACT 536
 Db 490 CGCAGCCTGTGTGATGATGTGTCAGCTACCGGACCTTACTTTCAGTGCCTATGACCAACT 549
 QY 537 GAAGGCTTCTGTGTGTCAGGACCTGACCTCTGACCTCTAGCAGCACCATGTTGGTGGTGG 596
 Db 550 GAAGGCTTCTGTGTGTCAGGACCTGACCTCTGACCTCTAGCAGCACCATGTTGGTGGTGG 609
 QY 597 CGCGTGGCGGCTGGGACCGCTGACCTGTGATCAGCCCGCTGGAGCTTATGCGGAGAAA 656
 Db 610 CGCGTGGCGGCTGGGACCGCTGACCTGTGATCAGCCCGCTGGAGCTTATGCGGAGAAA 669
 QY 657 GCTGACGCTCAGCATGTGTCTACCGGAGCTGGGTGCTGTGTTCGAACTTCGAGTGGC 716
 Db 670 GCTGACGCTCAGCATGTGTCTACCGGAGCTGGGTGCTGTGTTCGAACTTCGAGTGGC 729
 QY 717 TCAGGCTGGCTGGCGCTCACTGTGCTGGGCTGGGCGGCGGCGGCGGCGGCGGCGG 776
 Db 730 TCAGGCTGGCTGGCGCTCACTGTGCTGGGCTGGGCGGCGGCGGCGGCGGCGGCGG 789
 QY 777 CTTCTAGCCCTGTACTGTTCATATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAG 836
 Db 790 CTTCTAGCCCTGTACTGTTCATATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAG 849
 QY 837 GCCGAGGACAGACTTCTGTGGGCTGAGCTTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 896
 Db 850 GCCGAGGACAGACTTCTGTGGGCTGAGCTTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 909
 QY 897 GCGTGCAGTGTGACTTCTACCTTTGACGTGTGAAGAGCTGGCTCAATGGTTCAG 956
 Db 910 GCGTGCAGTGTGACTTCTACCTTTGACGTGTGAAGAGCTGGCTCAATGGTTCAG 969
 QY 957 AGCGATGGAGGTGTGAGAGTGAACCCCTGTGATGTGAGCTCCACCTGGCTGTGGTGG 1016
 Db 970 AGCGATGGAGGTGTGAGAGTGAACCCCTGTGATGTGAGCTCCACCTGGCTGTGGTGG 1029
 QY 1017 GAGGATCCCGGCGGAGTGGGACCAAGAGGACTTTCGAGGCTTCTCTCGGATCAT 1076
 Db 1030 GAGGATCCCGGCGGAGTGGGACCAAGAGGACTTTCGAGGCTTCTCTCGGATCAT 1089
 QY 1077 CAAGGCTGCCCCCTCTCTGTGCTATCATGACACCTATGAGTTCGGGAAAAGCTTCTT 1136
 Db 1090 CAAGGCTGCCCCCTCTCTGTGCTATCATGACACCTATGAGTTCGGGAAAAGCTTCTT 1149
 QY 1137 CGAGAGCTGAACAGAGACCGGCTTCTGGGCGGCTGAAGAGGCGAAGAGGAGGAGCC 1196
 Db 1150 CGAGAGCTGAACAGAGACCGGCTTCTGGGCGGCTGAAGAGGCGAAGAGGAGGAGCC 1209
 QY 1197 CGTCTCTCCAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1256
 Db 1210 CGTCTCTCCAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1269
 QY 1257 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1316
 Db 1270 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1329
 QY 1317 CCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1376
 Db 1330 CCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1389
 QY 1377 GATCATCACTTACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1436
 Db 1390 GATCATCACTTACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1449
 QY 1437 TTCCTGTGTTG 1496


```

Db 1450 TTCCCTGTTGTTGTTAGTGGGATGTTCTCCAGGAACCAAGAGCCCTCAGCCTGGTG 1509
QY 1497 TAGTCTCCCTGACCCCTGTTAAATCTTAAAGTCTAAAGATGATGAATCTTC 1546
Db 1510 TAGTCTCCCTGACCCCTGTTAAATCTTAAAGTCTAAAGATGATGAATCTTC 1559

RESULT 4
BC009330
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1560)
Strausberg,R.
Direct Submission
Submitted (12-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dierich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 26 Row: 9 Column: 12.
F RES Location/Qualifiers
source 1. .1560
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:16684 IMAGE:4126268"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC_19"
/lab_host="DH10B-R"
/notes="Vector: POTB7"
103..1158
/codon_start=1
/product="Unknown (protein for MGC:16684)"
/protein_id="AAH09330.1"
/db_xref="GI:14424606"
/translation="MADQFAGISPLQQWVAGTGVVTSFMTPLDVKVRLQSRP
SVASLMSRSLWSLYKWKCLLYCNGVLELYLCPNGARCATAWFOFPRFTGMDA
FVKI VRHGTITLWSGLPATLVMTVPATAYFTADQLKAFICGRALTSLDLPAPVAG
ALARGTWTVI SPLMKRTKLOQHVSYRELCAVRTAVAGGWSRLWLGFTALRD
VPFSALYWFNVELVKMLNGIRKQDTSVGMSPFVAGISGIVAAVUTLPPDVFVQTRQ
VALGAEAVRVNPLVDSTWLLRRFRAESGKGLFAGFLPRIKAAPSCA.MISTYE
FKSFFQLNDRLIG"
BASE COUNT 296 a 488 g 323 t
ORIGIN
Query Match 96.9%; Score 1498; DB 9; Length 1560;

```

```

Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 33 AGGCGCGGACGACGTAATAGGCGCCGCTGGCCCTGGGCGCGCGCGACGACGAC 92
Db 18 AGGCGCGGACGCGGTATGGGCGCCGCTGGCCCTGGGCGCGCGCGACGACGAC 77

QY 93 CTAGAGCCGAGACTGAAGCTTCAAGATGGCTGAACAGGACCCCTGCGGGCATCAGCCCCCT 152
Db 78 CTAGAGCCGAGACTGAAGCTTCAAGATGGCTGAACAGGACCCCTGCGGGCATCAGCCCCCT 137

QY 153 CCAGCAAAATGCTGGCTCAGGACCGGGCTGTGTTACCTCTCTTTCATGACACCCCT 212
Db 138 CCAGCAAAATGCTGGCTCAGGACCGGGCTGTGTTACCTCTCTTTCATGACACCCCT 197

QY 213 GGAGCTGGTGAAGTTGGCTCTCAGCTCTCAGCGGCCCTCCATGCGCCAGGACGATCATGCC 272
Db 198 GGAGCTGGTGAAGTTGGCTCTCAGCTCTCAGCGGCCCTCCATGCGCCAGGACGATCATGCC 257

QY 273 TTCCTCCAGACTGTGGAGCTCTCCTATACCAATGGAAGTGCCTCCTGATTGCAATGG 332
Db 258 TTCCTCCAGACTGTGGAGCTCTCCTATACCAATGGAAGTGCCTCCTGATTGCAATGG 317

QY 333 TGTCTCTGAGCCCTCTGTACTGTGCGCCAAATGCTGCCCGCTGTGCCACTGTGTTTCAAGA 392
Db 318 TGTCTCTGAGCCCTCTGTACTGTGCGCCAAATGCTGCCCGCTGTGCCACTGTGTTTCAAGA 377

QY 393 CCTTACCCGCTTCACTGGCACCATGGATGCTTCGTGAAGATCGTGAGGACGAGGCGAC 452
Db 378 CCTTACCCGCTTCACTGGCACCATGGATGCTTCGTGAAGATCGTGAGGACGAGGCGAC 437

QY 453 CAGACCCCTCTGAGCGGCTCTCCCGCCACCTCTGGTGTGATGACTGTGCCAGTACCGGCAT 512
Db 438 CAGACCCCTCTGAGCGGCTCTCCCGCCACCTCTGGTGTGATGACTGTGCCAGTACCGGCAT 497

QY 513 CTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGTGTCAGCCCTGACCTCTGA 572
Db 498 CTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGTGTCAGCCCTGACCTCTGA 557

QY 573 CCTCTAGCACCCATGTTGGTGTGCGCGCTGGCGCCCTGGGCGACCGTGACTGTGATCAG 632
Db 558 CCTCTAGCACCCATGTTGGTGTGCGCGCTGGCGCCCTGGGCGACCGTGACTGTGATCAG 617

QY 633 CCCCCCTGGAGCTTATGCGGACAAAGTGTGAGGCTCAGCATGTGTGTCACCGGAGCTGGG 692
Db 618 CCCCCCTGGAGCTTATGCGGACAAAGTGTGAGGCTCAGCATGTGTGTCACCGGAGCTGGG 677

QY 693 TGCTGTGTTGAACTGACGTGGCTCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGG 752
Db 678 TGCTGTGTTGAACTGACGTGGCTCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGG 737

QY 753 CCCCCCTGGCTTCAGATGTGCCCTTCTCAGCCCTGTACTGGTTCACCTATCAGCTGGT 812
Db 738 CCCCCCTGGCTTCAGATGTGCCCTTCTCAGCCCTGTACTGGTTCACCTATCAGCTGGT 797

QY 813 GAGAGCTGGCTCAATGGTTCAGGCGGAGGACAGACTTCTGTGGGCTAGGCTGTGT 872
Db 798 GAGAGCTGGCTCAATGGTTCAGGCGGAGGACAGACTTCTGTGGGCTAGGCTGTGT 857

QY 873 GGCTGTGGCTTCACAGGAGCGGTGGCTGCACTGTGCTTACCTTTGACGTTGATAA 932
Db 858 GGCTGTGGCTTCACAGGAGCGGTGGCTGCACTGTGCTTACCTTTGACGTTGATAA 917

QY 933 GACCAACGCGAGCTGCTCTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCACT 992
Db 918 GACCAACGCGAGCTGCTCTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCACT 977

QY 993 GGACTTCACTGGCTGCTGCTGGGAGGATCCGGGCGGAGTCCGGGCGACCAAGGAGCTCT 1052
Db 978 GGACTTCACTGGCTGCTGCTGGGAGGATCCGGGCGGAGTCCGGGCGACCAAGGAGCTCT 1037

QY 1053 TGCAGGCTTCTTCTTCGATCATCAAGGCTGCCCCCTCTCTGTGCCATCATGATCAGAC 1112

```


Qy	1394	CCCCAGTTCAAGACCAAAATCTTCCAGCTGCCCCCTTGGTGTTCCTGTTGCTGCT	1453
Db	1394	CCCCAAGTTCAAGACCAAAATCTTCCAGCTGCCCCCTTGGTGTTCCTGTTGCTGCT	1453
Qy	1454	AGCTGGGCATGTCTCCAGGAACCAAGAACCCCTCAGCCTGGTGTAAGTCCCTGACCCCT	1513
Db	1454	AGCTGGGCATGTCTCCAGGAACCAAGAACCCCTCAGCCTGGTGTAAGTCCCTGACCCCT	1513
Qy	1514	GTTAATTCCTTAAGTCTCAAGATGATGAA	1542
Di	1514	GTTAATTCCTTAAGTCTCAAGATGATGAA	1542

RESULT 7	LOCUS	DEFINITION	SEQUENCE	LENGTH	DATE
AR116038	AR116038	Sequence 68 from patent US 6132973.	1643 bp	linear	PAT 16-MAY-2001
ACCESSION	AR116038	AR116038			
VERSION	AR116038.1	GI:14096360			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1643)				
TITLE	Lal,P., Hillman,J.L., Bandman,O., Shah,P., Au-Young,J., Yue,H.,				
JOURNAL	Guebler,K.J. and Corley,N.C.				
FEATURES	Human regulatory molecules				
source	Patent: US 6132973-A 68 17-OCT-2000;				
	Location/Qualifiers				
	1..1643				
	/organism="unknown"				
BASE COUNT	307 a 501 c 476 g 359 t				
ORIGIN					
Query Match	93.48%; Score 1443.8; DB 6; Length 1643;				
Best Local Similarity	99.9%; Pred. No. 2.8e-308;				
Matches 1445; Conservative	0; Mismatches 2; Indels 0; Gaps 0;				
QY 100	CAGGACTCAAGCTTCAGATGCTGACACGAGCCCTCGGGGATCAGCCCTTCAGCAA 159				
Db 187	CAGGACTCAAGCTTCAGG GCTGACGAGGACCTCGGGGATCAGCCCTTCAGCAA 246				
QY 160	ATGTTGGCTCAGGCAACCGGGCTGTGGTACCTCTCTCTTCATGACACCCCTGCAGCTG 219				
Db 247	ATGTTGGCTCAGGCAACCGGGCTGTGGTACCTCTCTCTTCATGACACCCCTGCAGCTG 306				
QY 220	GTGAAGTTGCCTGCAGTCTCAGGGCCCTCCATGGCCAGCGAGTGTGCTTCCTCC 279				
Db 307	GTGAAGTTGCCTGCAGTCTCAGGGCCCTCCATGGCCAGCGAGTGTGCTTCCTCC 366				
QY 280	AGACTGTGGAGCTCTCTCTATACCAATGGAAGTGCCTCTGTATGCAATGGTGTCTG 339				
Db 367	AGACTGTGGAGCTCTCTCTATACCAATGGAAGTGCCTCTGTATGCAATGGTGTCTG 426				
QY 340	GAGCCTCTGTACTGTGGCCAAATGTTGCCCGCTGTGCCACCTGGTTTCAAGACCCCTACC 399				
Db 427	GAGCCTCTGTACTGTGGCCAAATGTTGCCCGCTGTGCCACCTGGTTTCAAGACCCCTACC 486				
QY 400	CGCTTCATCGCACCATGGATGCCTCTGTGAAGATCGTGAGGCACGAGGCAACAGGACC 459				
Db 487	CGCTTCATCGCACCATGGATGCCTCTGTGAAGATCGTGAGGCACGAGGCAACAGGACC 546				
QY 460	CTCTGGAGCGGCTCCCGCCACCTGGTGATGACCTGTGCAGCTACCGCATCTATCTC 519				
Db 547	CTCTGGAGCGGCTCCCGCCACCTGGTGATGACCTGTGCAGCTACCGCATCTATCTC 606				
QY 520	ACTGCCATGACAACTGAAGGCTCTTCTGTGTGTGGCCCTGAGCCCTGACCTCTGACCTCTAC 579				
Db 607	ACTGCCATGACAACTGAAGGCTCTTCTGTGTGTGGCCCTGAGCCCTGACCTCTGACCTCTAC 666				
QY 580	GCACCCATGGTGGCTGGCGCGCTGGGCCCTGGGCAACCGTGTGATCAGGCCCTG 639				

Db 667 GCACCCATGTGCTGGCTGGCCGCTGGCCCGCTTGGGACACCGCTGACTGTGATCAGCCCCCTG 726
 QY 640 GAGCTTATGCGCAAAAGCTGACGGCTCAGCATGTCTGCTACCGGAGCTGGCTGCTGT 699
 Db 727 GAGCTTATGCGCAAAAGCTGACGGCTCAGCATGTGCTGACCGGAGCTGGCTGCTGT 786
 QY 700 GTTCGACTGACGTGCTCAGGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
 Db 787 GTTCGAACTGACGTGCTCAGGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
 QY 760 GCCCTTCAGATGTGCTTCTCAGCCCTGTACTGCTCACTGCTCACTGCTGCTGCTGCTGCT 819
 Db 847 GCCCTTCAGATGTGCTTCTCAGCCCTGTACTGCTCACTGCTGCTGCTGCTGCTGCTGCT 906
 QY 820 TGGCTCAATGGTTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 879
 Db 907 TGGCTCAATGGTTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 966
 QY 880 GGCACTCTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 939
 Db 967 GGCACTCTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1026
 QY 940 CGCCAGTCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 999
 Db 1027 CGCCAGTCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1086
 QY 1000 ACCTGCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1059
 Db 1087 ACCTGCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1146
 QY 1060 TTCTTCTTCCGATCATCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1119
 Db 1147 TTCTTCTTCCGATCATCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1206
 QY 1120 TTCCGCAAAAGCTTCTTCCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1179
 Db 1207 TTCCGCAAAAGCTTCTTCCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1266
 QY 1180 CAAGGAGCAGGAGCCCGCTCTTCCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1239
 Db 1267 CAAGGAGCAGGAGCCCGCTCTTCCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1326
 QY 1240 AAGTGCTTTTCTCAGCATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1299
 Db 1327 AAGTGCTTTTCTCAGCATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1386
 QY 1300 TTCCAGGCTGAGGCT 1359
 Db 1387 TTCCAGGCTGAGGCT 1446
 QY 1360 GCTGCTCAGTCTGCT 1419
 Db 1447 GCTGCTCAGTCTGCT 1506
 QY 1420 AGCT 1479
 Db 1507 AGCT 1566
 QY 1480 AAGGCTCAGCT 1539
 Db 1567 AAGGCTCAGCT 1626
 QY 1540 GAACCTTC 1546
 Db 1627 GAACCTTC 1633

RESULT 8

AF119864

ID AF119864

AC AF119864 standard; RNA; HUM; 1402 BP.

XX AF119864;

XX

SV AF119864.1
 XX 12-MAY-2000 (Rel. 63, Created)
 DT 09-MAY-2001 (Rel. 67, Last updated, Version 2)
 XX
 DE Homo sapiens PRO2163 mRNA, complete cds.
 XX
 KW
 XX
 XX Homo sapiens (human)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 XX
 RN [1]
 RP 1-1402
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
 RT "Functional prediction of the coding sequences of 79 new genes deduced by
 RL analysis of cDNA clones from human fetal liver";
 XX Unpublished.
 RN [2]
 RP 1-1402
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
 RT Submitted (13-JAN-1999) to the EMBL/GenBank/DBJ databases.
 RL Department of Experimental Hematology, Institute of Radiation Medicine,
 XX Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
 DR SWISS-PROT; Q9B2J4; CG69_HUMAN.
 XX
 FH Key Location/Qualifiers
 FT source 1..1402
 FT /db_xref="taxon:9606"
 FT /organism="Homo sapiens"
 FT /clone="PLB8041"
 FT /tissue_type="liver"
 FT /dev_stage="fetus"
 FT 4..1017
 FT /codon_start=1
 FT /db_xref="SWISS-PROT:Q9B2J4"
 FT /note="Predicted protein of HQ2163"
 FT /product="PRO2163"
 FT /protein_id="AAP69618.1"
 FT /translation="MVASGTGAVTSLFMTPLDVVKVRLQSORNSASELMPSSRLMSL
 SYTKWKLJYCMGVLEPLYLCPNGARCATWFQDPTFTGTMDFVKI VRHGTGTLMSG
 LPATLVMTVPATAIYFTAYDOLKAPLGRALTSDLYAPVAGALALCTVTYISPLELM
 RTKLOAHVSVYRELGACVRTAVAGGWSLWLGWPTALRDVPPSALYFNFYET VKSNL
 NGLRPKDOTSVGWSFVAGGISRTVAVALTLFPDVVKIQRQVAGLAWAVRNP.LHVDST
 WLLLRIRAESGTGLFAGFLPRLIIPAAPSALIMISTYEFKSPFQRLNQRLIGG"
 XX
 SQ Sequence 1402 BP; 261 A; 432 C; 401 G; 308 T; 0 other;
 Query Match 89.3%; Score 1381.2; DB 17; Length 1402;
 Best Local Similarity 99.8%; Pred. No. 1.9e-294;
 Matches 1383; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 157 CAAATGGTGGCTCAGCACCGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGGAC 216
 Db 1 CAAATGGTGGCTCAGCACCGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGGAC 60
 QY 217 GTGGTGAAGGTTGCTGAGTCTCAGCGGCCCTCCATGCGCAGAGTGTGCTTCC 276
 Db 61 GTGGTGAAGGTTGCTGAGTCTCAGCGGCCCTCCATGCGCAGAGTGTGCTTCC 120
 QY 277 TCCAGACTGTGGAGCTCTCTTATACCAATGGAGTGCCTCTCTGTTGCAATGGTCTC 336
 Db 121 TCCAGACTGTGGAGCTCTCTTATACCAATGGAGTGCCTCTCTGTTGCAATGGTCTC 180
 QY 337 CTGGAGCTCTGTACTGTCGCCAAATGGTGCCTGTCGCCACCTGGTTTCAAGACCT 396
 Db 181 CTGGAGCTCTGTACTGTCGCCAAATGGTGCCTGTCGCCACCTGGTTTCAAGACCT 240

QY 397 ACCCGCTTCACTGGCAACCATGATGCTTGTGTAAGATCGTGAGGACGAGGCGACACAGG 456
Db 241 ACCCGCTTCACTGGCAACCATGATGCTTGTGTAAGATCGTGAGGACGAGGCGACACAGG 300
QY 457 ACCCTGTGGAGGGGCTCCCGCCACACCTGGTGATGACTGTGCCAGCTACCGCCATCTAC 516
Db 301 ACCCTGTGGAGGGGCTCCCGCCACACCTGGTGATGACTGTGCCAGCTACCGCCATCTAC 360
QY 517 TTCACTGCTATGACCACTGAGGCTTCTGTGTGTCGAGCCCTGACCTCTGACCTC 576
Db 361 TTCACTGCTATGACCACTGAGGCTTCTGTGTGTCGAGCCCTGACCTCTGACCTC 420
QY 577 TAGCGACCATGCTGGCGGCTGCCCGCTGGGCAACCGTGGGCACTGTGATCAGCC 636
Db 421 TAGCGACCATGCTGGCGGCTGCCCGCTGGGCAACCGTGGGCACTGTGATCAGCC 480
QY 637 CTGAGCTTATCGGACAAAGCTCGAGGCTCAGCATGTGTGTACCGGAGCTGGGTGCC 696
Db 481 CTGAGCTTATCGGACAAAGCTCGAGGCTCAGCATGTGTGTACCGGAGCTGGGTGCC 540
QY 697 TGTGTTGCAACTGCACTGCTCAGGCTGGGCTGCGCTCACTGTGGCTGGGCTGGGCTGG 756
Db 541 TGTGTTGCAACTGCACTGCTCAGGCTGGGCTGCGCTCACTGTGGCTGGGCTGGGCTGG 600
QY 757 ACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCACTATGAGCTGGTGAAG 816
Db 601 ACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCACTATGAGCTGGTGAAG 660
QY 817 AGCTGGCTCAATGGGTTGAGCGAAGGACCAAGATCTGTGTGGCATGAGCTTTGTGGCT 876
Db 661 AGCTGGCTCAATGGGTTGAGCGAAGGACCAAGATCTGTGTGGCATGAGCTTTGTGGCT 720
QY 877 GGTGGCTATCAGGACGCTGGCTCAGGCTGGGCTGCGCTCACTGTGGCTGGGCTGGGCTGG 936
Db 721 GGTGGCTATCAGGACGCTGGCTCAGGCTGGGCTGCGCTCACTGTGGCTGGGCTGGGCTGG 780
QY 937 CAACGCCAGTCTGCTGGGACGATGGAGGCTGTGAGATGAACCCCTGCATGTGAC 996
Db 781 CAACGCCAGTCTGCTGGGACGATGGAGGCTGTGAGATGAACCCCTGCATGTGAC 840
QY 997 TCCACCTGCTGCTGCTGGGAGGATCGCGGCGGAGTCGGGCGGAGTCGGGCAACAGGAGCTTTGCA 1056
Db 841 TCCACCTGCTGCTGCTGGGAGGATCGCGGCGGAGTCGGGCGGAGTCGGGCAACAGGAGCTTTGCA 900
QY 1057 GGTCTCTTCTCGGATCATCAGGCTGCCCTTCTGCGCATCATGATCAGCACTAT 1116
Db 901 GGTCTCTTCTCGGATCATCAGGCTGCCCTTCTGCGCATCATGATCAGCACTAT 960
QY 1117 GAGTTGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTTCTGGGCGGCTGAAG 1176
Db 961 GAGTTGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTTCTGGGCGGCTGAAG 1020
QY 1177 GGGCAAGGAGCAAGGACCGGCTTCTTCCAGATGGGAGGAGGAGGAGGAGCAACCA 1236
Db 1021 GGGCAAGGAGCAAGGACCGGCTTCTTCCAGATGGGAGGAGGAGGAGGAGGAGCAACCA 1080
QY 1237 GCGAAGTGGCTTTTCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1296
Db 1081 GCGAAGTGGCTTTTCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1297 AAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1356
Db 1141 AAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1357 CTGCTGCTCAGTCTGGGATCATCACTTACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1416
Db 1201 CTGCTGCTCAGTCTGGGATCATCACTTACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1417 TCCAGCTGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1476
Db 1261 TCCAGCTGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1477 AAGAAGCCCTCAGCCTGGTGTAGTCTCCCTGACCCCTTTGTTAAATTCCTTAAGTCTAAGAT 1536

Db 1321 AAGAAGCCCTCAGCCTGGTGTAGTCTCCCTGACCCCTGTAAATTCCTTAAGTCTAAGAT 1380
QY 1537 GATGAA 1542
Db 1381 GATGAA 1386
RESULT 9
AX403084
LOCUS AX403084 1114 bp DNA linear PAT 07-JUN-2002
DEFINITION Sequence 1 from Patent WO0198355.
ACCESSION AX403084
VERSION AX403084.1 GI:21388033
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Lewin, D., Adams, S.H. and Yu, X.X.
Cgi-69 compositions and methods of use
Patent: WO 0198355-A 1 27-DEC-2001;
Genentech, Inc. (US); Curagen Corporation (US)
FEATURES
source
1..1114
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 194 a 349 c 332 g 239 t
ORIGIN
Query Match 68.2%; Score 1054.4; DB 6; Length 1114;
Best Local Similarity 97.8%; Pred. No. 2,6e-222;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 105 CTCAGCTTCAAGATGGCTGACAGGACCCCTCGGGCATACGCCCTCCAGCAAAATGGT 164
Db 1 CTCAGCTTCAAGATGGCTGACAGGACCCCTCGGGCATACGCCCTCCAGCAAAATGGT 60
QY 165 GGCCTCAGGACACCGGGCTGTGGTACCTCTCTTCATGACACCCCTGAGCGTGAA 224
Db 61 GGCCTCAGGACACCGGGCTGTGGTACCTCTCTTCATGACACCCCTGAGCGTGAA 120
QY 225 GGTTCGCTGAGTCTCAGCGGCCCTCCATGCCAGGAGCTGATGCTTCCTCCGACT 284
Db 121 GGTTCGCTGAGTCTCAGCGGCCCTCCATGCCAGGAGCTGATGCTTCCTCCGACT 180
QY 285 GTGGAGCTCTCTTACCAAT-----GGAAGTGCCTCT 320
Db 181 GTGGAGCTCTCTTACCAATTTGCCCTCTCTTCCAAATGCCAGGAGTGCCTCT 240
QY 321 GTATTGCAATGGTCTCTGAGGCTCTGTACCTGTGCCAAATGGTCCCGCTGTGCCAC 380
Db 241 GTATTGCAATGGTCTCTGAGGCTCTGTACCTGTGCCAAATGGTCCCGCTGTGCCAC 300
QY 381 CTGGTTTCAAGACCTTACCGCTTCACTGGCACCATGATGCTTCGTGAAGATCTGAG 440
Db 301 CTGGTTTCAAGACCTTACCGCTTCACTGGCACCATGATGCTTCGTGAAGATCTGAG 360
QY 441 GCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 500
Db 361 GCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 501 AGCTACCCCATCTACTTCACTGCTTATGACCAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 560
Db 421 AGCTACCCCATCTACTTCACTGCTTATGACCAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 561 CTGAGCTCTGAGCTTACGACCAATGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 620
Db 481 CTGAGCTCTGAGCTTACGACCAATGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 540
QY 621 GACTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 680


```

Db 541 GACTGTGATCAGCCCCCTGGAGCTTATGCGGACAAAGCTGCGAGGCTCAGCATGTGCTGA 600
QY 681 CCGGAGAGTGGTGGCTGTGTTCGAACATGAGTGGCTCAGGTTGGCTGGCTCAGCTCAGTGTG 740
Db 601 CCGGAGAGTGGTGGCTGTGTTCGAACATGAGTGGCTCAGGTTGGCTGGCTCAGCTCAGTGTG 660
QY 741 GCTGGGCTGGGCCCCACCTGACCTTCAGATGTGCTCCTTCTCAGCCCTGTACTGGTTCAA 800
Db 661 GCTGGGCTGGGCCCCACCTGACCTTCAGATGTGCTCCTTCTCAGCCCTGTACTGGTTCAA 720
QY 801 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGGCGGAGGACCAAGAGCTGTGTGG 860
Db 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGGCGGAGGACCAAGAGCTGTGTGG 780
QY 861 CATGAGCTTTGGTGGTGGCTATCTCAGGAGCGGTGGCTGAGTGTGACTCTACCCCTT 920
Db 781 CATGAGCTTTGGTGGTGGCTATCTCAGGAGCGGTGGCTGAGTGTGACTCTACCCCTT 840
QY 921 TGACGTGTAAAGACCCCAACCCAGTGGCTTCGGAGGAGTGGAGGCTGTGAGAGTGAA 980
Db 341 TGACGTGTAAAGACCCCAACCCAGTGGCTTCGGAGGAGTGGAGGCTGTGAGAGTGAA 900
QY 981 CCCCCTGATGTGAGTCCCACTGGCTGTGCTGCGGAGGATCCGGGCGGAGTGGGCGAC 1040
Db 901 CCCCCTGATGTGAGTCCCACTGGCTGTGCTGCGGAGGATCCGGGCGGAGTGGGCGAC 960
QY 1041 CAAGGAGCTTTGGAGGCTTCCTTCTGATCATCAAGGCTGCCCCCTCTGTGCCAT 1100
Db 961 CAAGGAGCTTTGGAGGCTTCCTTCTGATCATCAAGGCTGCCCCCTCTGTGCCAT 1020
QY 1101 CATGATCAGCACTATGATTCGGCAAGAGCTTCTTCAGAGGCTGAACAGGAACGGCT 1160
Db 1021 CATGATCAGCACTATGATTCGGCAAGAGCTTCTTCAGAGGCTGAACAGGAACGGCT 1080
QY 1161 TTGGGCGGCTGAAGGGGCAAGAGGCGAAGGAC 1194
Db 1081 TTGGGCGGCTGAAGGGGCAAGAGGCGAAGGAC 1114

RESULT 10
AF317711
LOCUS Homo sapiens mitochondrial carrier protein CGI-69 long form mRNA,
DEFINITION complete cds.
ACCESSION AF317711
VERSION AF317711.1 GI:12620399
KEYWORDS
SC "E
NISM
REFERENCE
AUTHORS Yu,X.X., Lewin,D.A., Zhong,A., Brush,J., Schow,P.W., Sherwood,S.W.,
Pan,G. and Adams,S.H.
TITLE Overexpression of the human 2-oxoglutarate carrier lowers
mitochondrial membrane potential in HEK-293 cells: contrast with
the unique cold-induced mitochondrial carrier CGI-69
JOURNAL Biochem. J. 353 (Pt 2), 369-375 (2001)
MEDLINE 2106063
PUBMED 11139402
REFERENCE
AUTHORS Yu,X.X., Lewin,D.A., Zhong,A., Brush,J., Schow,P.W., Sherwood,S.W.,
Pan,G. and Adams,S.H.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2000) Endocrinology, Genentech, Inc., 1 DNA Way,
South San Francisco, CA 94080, USA
FEATURES
source
1. .1114
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue type="liver"
14. .1033
/note="similar to CGI-69 encoded by GenBank Accession

```

```

Number AF151827"
/codon_start=1
/product="mitochondrial carrier protein CGI-69 long form"
/protein_id="AAG60687.1"
/db_xref="GI:12620400"
/translation="MADQDPAGISPLQQWVASGTGAVVTSLEWTPLDVVKVRLQSORP
SMASLMPSSRLMSLYTKLPSLSASTGKLLYXNGVLEPLICPNAGKATWFDQPT
RTGTMDAFVKIVRHEGRTMLSGPLATVMTATATYFYTDQKALICGRALTSQ
LYAPWAGALARGITVVISPLELAKTQAHVSYRELCACTVATVAGGWSLWLG
DVTALRDVPFSAIYMFNVELVKSWLNGLRPKDQTSVMSGFVAGGSGTVAVALPLPF
DVVKTQROVALGAMEARVNPPLHVDSTWLLLRIRAESGKGLFAGFLPRIIKAAPSC
AIMISTVEFGKSFQRLNQDRLLGG"
BASE COUNT 194 a 349 c 332 g 239 t
ORIGIN
Query Match 68.2%; Score 1054.4; DB 9; Length 1114;
Best Local Similarity 97.8%; Pred. No. 2.6e-222;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 105 CTGAGGCTTCAGATGGCTGACAGGACCCCTGGGCGCATGCCCTTCCAGCAAAATGGT 164
Db 1 CTGAGGCTTCAGATGGCTGACAGGACCCCTGGGCGCATGCCCTTCCAGCAAAATGGT 60
QY 165 GGCCTCAGGACACCGGGGCTGTGTACTCTCTCTCTATGACACCCCTGGAGCTGTGAA 224
Db 61 GGCCTCAGGACACCGGGGCTGTGTACTCTCTCTCTATGACACCCCTGGAGCTGTGAA 120
QY 225 GGTTCGCTGCGAGTCTCAGCGGCTCTCATGGCCAGCGAGTGTGCTTCTCCAGACT 284
Db 121 GGTTCGCTGCGAGTCTCAGCGGCTCTCATGGCCAGCGAGTGTGCTTCTCCAGACT 180
QY 285 GTGAGGCTCTCTATACAAAT-----GGAGTGGCTCT 320
Db 181 GTGAGGCTCTCTATACAAATTTGCCCTCTCTCTCCAAATCCACAGGAAGTGGCTCT 240
QY 321 GTATTGCAATGTGTCTGGAGCTCTGTACTGTGCCCAATGTGCCCAATGTGCCCTGTGCCAC 380
Db 241 GTATTGCAATGTGTCTGGAGCTCTGTACTGTGCCCAATGTGCCCAATGTGCCCTGTGCCAC 300
QY 381 CTGTTTCAAGACCTACCGCTCTACTGGCCACCATGGATGCTTCTGTGAAGATCTGTAG 440
Db 301 CTGTTTCAAGACCTACCGCTCTACTGGCCACCATGGATGCTTCTGTGAAGATCTGTAG 360
QY 441 GCACGAGGCGACCAAGGACCTCTGGAGCGGCTTCCGCCCAATGTGCCCTGTGCCCTGTGCC 500
Db 361 GCACGAGGCGACCAAGGACCTCTGGAGCGGCTTCCGCCCAATGTGCCCTGTGCCCTGTGCC 420
QY 501 AGCTACCGCCATCTACTCTACTGCTATGACCACTGAGGCTTCTCTGTGTGGTGGAGC 560
Db 421 AGCTACCGCCATCTACTCTACTGCTATGACCACTGAGGCTTCTCTGTGTGGTGGAGC 480
QY 561 CTGACCTCTGACCTCTACGCCACCATGGTGGCGCTGGCGGCTGGGACCGT 620
Db 481 CCGTACCTCTGACCTCTACGCCACCATGGTGGCGCTGGCGGCTGGGACCGT 540
QY 621 GACTGTGATCAGCCCTCTGAGCTTATGGGACAAAGCTGAGGCTCAGAGCTGTGCTGA 680
Db 541 GACTGTGATCAGCCCTCTGAGCTTATGGGACAAAGCTGAGGCTCAGAGCTGTGCTGA 600
QY 681 CCGGAGCTGGTGGTGGTCTGTAAGTGTGAGTGGCTGAGGCTGGGCTGAGCTGTG 740
Db 601 CCGGAGCTGGTGGTGGTCTGTAAGTGTGAGTGGCTGAGGCTGGGCTGAGCTGTG 660
QY 741 GCTGGGCTGGGCCCCACCTGACCTTTCAGATGTGCTCCTTTCAGCCCTGTACTGTTCAA 800
Db 661 GCTGGGCTGGGCCCCACCTGACCTTTCAGATGTGCTCCTTTCAGCCCTGTACTGTTCAA 720
QY 801 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGGCGGAGGACCAAGAGCTGTGTGG 860
Db 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGGCGGAGGACCAAGAGCTGTGTGG 780
QY 861 CATGAGCTTTGGTGGTGGCTATCTCAGGAGCGGTGGCTGAGTGTGACTCTACCCCTT 920

```


QY 622 ACTGTGATCAGCCCTTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGCGTAC 681
 Db 73006 ACGGTGGTGCAGCCCTTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGCGTAC 73065
 QY 682 CGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 741
 Db 73066 CGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 73125
 QY 742 CTGGGCTGGGCGCCCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 801
 Db 73126 CTGGGCTGGGCGCCCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 73185
 QY 802 TATGAGCTGGTGAAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 861
 Db 73186 TATGAGCTGGTGAAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 73245
 QY 862 ATGAGCTTTGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 920
 Db 73246 ATGAGCTTTGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 73305
 QY 921 TGAGCTGGTAAAGACCAAGCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 980
 Db 73306 TGAGCTGGTAAAGACCAAGCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 73365
 QY 981 CCCCCTG ATGTGAGCTCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1040
 Db 73366 CCCCCTG ATGTGAGCTCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 73424
 QY 1041 CAGGGACTTTTGGAGCTTTCTTCTCGGATCATCAGGCTGCCCTCTCTCTCTCTCTCT 1100
 Db 73425 CAGGGACTTTTGGAGCTTTCTTCTCGGATCATCAGGCTGCCCTCTCTCTCTCTCTCT 73481
 QY 1101 CATGATCAGACTATGAGTGGCGAAAGCTTTCTTCTCGGAGCTGAAACAGGACCGGCT 1160
 Db 73482 CATGATCAGACTATGAGTGGCGAAAGCTTTCTTCTCGGAGCTGAAACAGGACCGGCT 73541
 QY 1161 TCTGGCGGCTGAAAG 1176
 Db 73542 TCTGGCGGCTGAAAG 73557

 RESULT 13
 AC111356
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-161G12, *** SEQUENCING IN PROGRESS
 A^c STON
 V^c ON
 KE^c JDS
 SOURCE Norway rat
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 156175)
 Murny D.M., Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C.,
 Albrooks S.L., Amaratunge H.C., Are J.R., Ayelle M., Banks T.,
 Barbara J., Benton J., Bimagne K., Blankenburg K., Bonnin D.,
 Bouck J., Bowie S., Brieva M., Brown E., Brown M., Bryant N.P.,
 Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
 Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,
 Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
 Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
 Davila M.L., Davis C., Davy-Carroll L., Dederich D.A.,
 Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,
 Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,
 Earnhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,
 Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Prantz P.,
 Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
 Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,
 Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,
 Hernandez C., Hodgson A., Hognes M., Holloway C., Hollins B.,

Homsí, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge O., Liu, C., Liu, J., Liu, W., Loulseg, H.,
 Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, S., Martindale, A., Martinez, E.,
 Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Meig, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, K.,
 Nguyen, N., Nickerson, E., Nwokkwo, S., Oguh, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
 Sodergren, E., Sotnik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, E., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 156175)
 Worley, K.C.
 Direct Submission
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 156175)
 Worley, K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:16701120.

 Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GMCE
 Center clone name: CH230-161G12

 Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 99973 bases at least Q40
 Consensus quality: 106615 bases at least Q30
 Consensus quality: 111100 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 66 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 1025: contig of 1025 bp in length
 1026 1125: gap of unknown length
 1126 2118: contig of 1093 bp in length
 2119 2318: gap of unknown length
 2319 3454: contig of 1136 bp in length
 3455 3554: gap of unknown length
 3555 4690: contig of 1136 bp in length
 4691 4790: gap of unknown length
 4791 5955: contig of 1165 bp in length

```

* 5956 6055: gap of unknown length
* 6056 7439: contig of 1384 bp in length
* 7440 7539: gap of unknown length
* 7540 8707: contig of 1168 bp in length
* 8708 8807: gap of unknown length
* 8808 10033: contig of 1226 bp in length
* 10034 11778: contig of 1645 bp in length
* 11779 13633: contig of 1755 bp in length
* 13634 14947: contig of 1214 bp in length
* 14948 15047: gap of unknown length
* 15048 16569: contig of 1522 bp in length
* 16570 17886: contig of 1117 bp in length
* 17887 18788: gap of unknown length
* 18789 19738: contig of 1852 bp in length
* 19739 21533: contig of 1694 bp in length
* 21534 22815: contig of 1083 bp in length
* 22816 24052: contig of 1237 bp in length
* 24053 24152: gap of unknown length
* 24153 25250: contig of 1098 bp in length
* 25251 25350: gap of unknown length
* 25351 26944: contig of 1594 bp in length
* 26945 28474: gap of unknown length
* 28475 28574: contig of 1430 bp in length
* 28575 29925: contig of 1351 bp in length
* 29926 30025: gap of unknown length
* 30026 31402: contig of 1377 bp in length
* 31403 31502: gap of unknown length
* 31503 33645: contig of 2143 bp in length
* 33646 33745: gap of unknown length
* 33746 35290: contig of 1545 bp in length
* 35291 35390: gap of unknown length
* 35391 36672: contig of 1282 bp in length
* 36673 38565: contig of 1793 bp in length
* 38566 40077: gap of unknown length
* 40078 40177: contig of 1412 bp in length
* 40178 41278: contig of 1101 bp in length
* 41279 42846: contig of 1468 bp in length
* 42847 42946: gap of unknown length
* 42947 44779: contig of 1833 bp in length
* 44780 44879: gap of unknown length
* 44880 47030: contig of 2151 bp in length
* 47031 47130: gap of unknown length
* 47131 48946: contig of 1716 bp in length
* 48947 50071: gap of unknown length
* 50072 50171: contig of 1125 bp in length
* 50172 51617: contig of 1445 bp in length
* 51618 53821: contig of 2104 bp in length
* 53822 55611: contig of 1690 bp in length
* 55612 55711: gap of unknown length
* 55712 57525: contig of 1814 bp in length
* 57526 59106: contig of 1481 bp in length
* 59107 59206: gap of unknown length
* 59207 61142: contig of 1936 bp in length
* 61143 64411: contig of 3169 bp in length
* 64412 66866: contig of 2175 bp in length
* 66867 66786: gap of unknown length

```

```

* 66787 68966: contig of 2180 bp in length
* 68967 72130: contig of 3064 bp in length
* 72131 72230: gap of unknown length
* 72231 73778: contig of 1548 bp in length
* 73779 76273: gap of unknown length
* 76274 76373: contig of 2395 bp in length
* 76374 78932: contig of 2559 bp in length
* 78933 79032: gap of unknown length
* 79033 81984: contig of 2951 bp in length
* 81984 82084: gap of unknown length
* 82084 84717: contig of 2633 bp in length
* 84717 87800: gap of unknown length
* 87801 87900: contig of 2984 bp in length
* 87901 92353: contig of 4453 bp in length
* 92354 92454: gap of unknown length
* 92454 96500: contig of 4046 bp in length
* 96500 96600: gap of unknown length
* 96600 99504: contig of 2904 bp in length
* 99504 99604: gap of unknown length
* 99604 103251: contig of 3647 bp in length
* 103251 106405: gap of unknown length
* 106405 106405: contig of 3055 bp in length

Query Match      34.7%   Score 537.2; DB 2; Length 156175;
Best Local Similarity 77.6%; Pred. No. 2.5e-108;
Matches 709; Conservative 0; Mismatches 183; Indels 22; Gaps 4;

QY 263 AGCTGATGCTTCCCTCCAGAGCTGGAGCTCTCTCTATACCAAAATGAAGTGCCTCTCT 322
Db 77061 AGATTCTGAGTCTCTCTCACTACCTAATCACTCTGTCTCTCGGGAAGGCTCTCCAT 77120

QY 323 ATTGCATGCTGCTGGAGCTCTGTACTGTGCCAAATGTCGCCGCTGTCACCT 382
Db 77121 ACTGCTATGCTGCTGGAGCTCTGGAGCTGTGCCAAATGATACCTGCTGCGCACCT 77180

QY 383 GGTTCACAGACCTACCGCTTCTACCTGGCACCACATGATGCTTCTGTAAGATCGTAGGC 442
Db 77181 GGTTCAGATCTTACAGGTTTCTGGCACCTTGGATGCTTGTGAAGATGTGAGGC 77240

QY 443 ACAGGGCACAGACCTCTGGAGGGGCTCTCCCGCCACCCTGGTGGTATGATGTCGAC 502
Db 77241 -CGAGGGCACAGACCTCTGGAGGGGCTCTCCAGCTACCTGATGATGATCTATCAG 77299

QY 503 CTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGTTGGTTCAGGCC 562
Db 77300 CTACCGCATCTATTTCATTTGCTTATGACCACTCAAG---ACCTATGCTCAATCCT 77355

QY 563 TGACCTCTGACCTCTACGACCCATGTTGGTGGGGCTGGCCGCCCTGGCACCGTGA 622
Db 77356 TGACCTCTGACCTCTACGACCCACCGTGGCTGGTGGCTGACCCGAATGGCACCGTGA 77415

QY 623 CTGTATCAGCCCTCTGAGCTTATGCGGACAAAGCTGACAGCTGATGTCGTACC 682
Db 77416 CAGTGTGAGCCCTCTGAGCTTGTGCGAACCAAGCTGACAGCTGATGTCATACC 77475

QY 683 GGGAGCTGGTGTGTTTCAACTGAGTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 742
Db 77476 GCAAACTGGCTTACTGTTTCAAACTGAGTGGCTGAGGCTGAGGCTGAGGCTGAGGCT 77535

QY 743 TGGGCTGGGGCCCACTGCTCTCGAGATGTCCTTCTCAGCCCTGACTGTTCACT 802
Db 77536 TGGGCTGGGCTTCAATAGTCTTCAAGATGACCCCTTGGCAGCTCTCCACTGGTCAACT 77595

QY 803 ATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTACGGCCGAGGACCACTTCTGTGGCA 862
Db 77596 ACAAGCTGGTGAAGGCGCAGCTGAAATGGGCGGACGACTGAAAGACCACTGAGCA 77655

QY 863 TGAGCTTGGTGTGCTGCTGATCTCAGGAGCGTGGCTGAGTGTGCTGACTTACCTTTG 922
Db 77656 TCAGTTTGTGCTTAATGGCATCTCAGGAATGGGGCTGCGCACCTTACCTTCTTTG 77715

```

QY 923 ACCTGCTTAAGACCCACGCGGCTGCTCTTGGAGGCGATGGAGGCTGTGAGAGTGAACC 982
 Db 77716 ATGTGATCAAGACACACGACAAATGTCACTGTGAGCGGTGAGGCTATGAGA ----- 77768
 QY 983 CCCTGCTATGGACTCCACCTGCTGCTGCTGGAGGATCGGGCGGATCGGGACCA 1042
 Db 77769 -----GTGCTCTCACCTGGCTCTGCTTGGAGAAATCCAGGCTGAATCTGGACCA 77820
 QY 1043 AGGACTCTTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCAACGAGCGGCTTC 1102
 Db 77821 GGGGACTC--TGCAGGTTTCTTCCCTCCAGGATCATCAAGGCTGCAACCTTCTGTGCCATCA 77878
 QY 1103 TGATCAGACCTATGATGTGGCAAAAGCTTCTTCCAGAGGCTGAACGAGCGGCTTC 1162
 Db 77879 GCAGCAGCTATGATGTGGCAAAAGCTTCTTCCAGAGGCTGAACGAGCGGCTTC 77938
 QY 1163 TGGCGGCTGAAAG 1176
 Db *939 TGGGCACTGAAAG 77952

AC123139 128802 bp DNA linear HTG 13-JUL-2002
 Rattus norvegicus clone CH230-11K3, *** SEQUENCING IN PROGRESS ***
 29 unordered pieces.
 AC123139
 AC123139.2 Gi:21686901
 HTG: HTGS PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 128802)

REFERENCE
AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Burch, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Jacobson, B., Jia, Y., Johnson, R., Jolly, S., Hume, J., Jackson, L.E., Karlsson, E., Kelly, S., Khan, U., King, J., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Loulsegged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, B., Massey, B., Maxwell, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savory, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,

1 1434: contig of 1434 bp in length
 1435 1534: gap of unknown length
 1535 3218: contig of 1684 bp in length
 3219 3318: gap of unknown length
 3319 4531: contig of 1213 bp in length
 4532 4631: gap of unknown length
 4632 5802: contig of 1171 bp in length
 5803 5902: gap of unknown length
 5903 7060: contig of 1158 bp in length
 7061 7160: gap of unknown length
 7161 8514: contig of 1354 bp in length
 8515 8614: gap of unknown length
 8615 9672: contig of 1058 bp in length
 9673 9773: gap of unknown length
 9773 12180: contig of 2408 bp in length
 12181 12280: gap of unknown length
 12281 13463: contig of 1183 bp in length
 13464 13563: gap of unknown length
 13564 16106: contig of 2543 bp in length
 16107 16206: gap of unknown length
 16207 18884: contig of 2678 bp in length
 18885 19884: gap of unknown length
 19885 20532: contig of 1548 bp in length
 20533 22633: gap of unknown length
 22634 23210: contig of 2578 bp in length
 23211 25515: contig of 2205 bp in length
 25516 25616: gap of unknown length
 25617 27411: contig of 1796 bp in length

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 29 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.
 Direct Submission
 2 (bases 1 to 128802)
 Worley, K.C.
 Direct Submission
 Submitted (29-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 128802)
 Worley, K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jul 3, 2002 this sequence version replaced gi:21239891.

COMMENT

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GMFG
 Center clone name: CH230-11K3
 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 108533 bases at least Q40
 Consensus quality: 111720 bases at least Q30
 Consensus quality: 113548 bases at least Q20

	Query Match	34.38%	Score 530	DB 2	Length 128802	
	Best Local Similarity	75.9%	Prod. No. 1e-106			
	Matches 733	Conservative	0	Mismatches 200	Indels 33	Gaps 5
QY	50	TGGGCGCGCTTGGCGCTGGCGCGCGCC	JCAGGAGACACGCTTAGGCGCAGGACTGAA	109		
Db	78125	TGGGCGCCACCGCGCCCTTGGCGTGGCGGG	AAGCAGCCGCTCAGAGCAGGATGAA	78066		
QY	110	GCTTCAGATGGCTGACCAAGGACCTCGGSGCAT	CAGCCCTCCAGCAAAATGCTGGCT	169		
Db	78065	ATTCCNAGATGGTTGACCAAGGATCTGGGTGCAT	TAAGACCTTTTACAAATGGTGGCT	78006		
QY	170	CAGGCACCGGGCTGTGTTACTCTCTCTCAT	TGACACCCCTGGAGCTGTCTCCAGACT	229		
Db	78005	CAGGAATGGGGTTGTGTACATCCCTCTCTCAT	GACACCCCTGTTGTAGTGAAGTCC	77946		
QY	230	GCCTGCAGTCTCA-----CGGCGCCCTC	CATGGCCAGGAGCTGATGCTTCTCCAGACT	284		
Db	77945	GCTTTCAGTCTCAAGACCTGTCTCTCAT	TAGTACAAATGTCACACCTCCAGATT	77886		
QY	285	GTGGAGGCTCTCTTATACAAAT-----	GGAAAGTGCTCCT	320		
Db	77885	CTGGAGTCTTCTCATACCAAAATCACCCCTCT	CTCAANTCCCACTGACATGCTTCT	77826		
QY	321	GTATTGCAATGGTCTTGAGGCTCTGTACT	TGTCGCAAAATGGTGCCGCTGTGCAC	380		
Db	77825	GTACTCAATGGTGTCTTGAGGCCCTGTACT	TGTCGCAAAATGGTACTCGCTGCACAC	77766		
QY	381	CTGTTTTCAGACCCCTACCCTTCACTGSCAC	TGATGCTCTTGTGAGATCGTAG	440		
Db	77765	CTGTTTCAGATCTTACAGATTTCACTAGAC	CTTGTAGTGCCTTTGTGAAGACTGTAG	77706		
QY	441	GCACGAGGCGACAGGACCTCTTGAGCGGCG	CTCCCGCCACCTGTGATCACTGTGCC	500		
Db	77705	GCATGAGGCGCACAGGACCTGTGAGAGAGC	CTCCAGCCACGTTGATGATGCTGTGTC	77646		
QY	501	AGCTACCGGCATCTACTTCACTGCCTATGAC	CAACTGAGGCCCTTCTGTGTGGTCGAGC	560		


```

Db 121 TGAACACGAGCGGCTTCTGGGCGGTGAAGGGGCAAGGAGCAAGGACCCCGTCTCTC 180
QY 1205 CCACGGATGGGAGAGGGGAGGAGAGACCCAGCCAAAGTGCCCTTTTCCTCAGCACTGAGG 1264
Db 181 CCACGGATGGGAGAGGGGAGGAGAGACCCAGCCAAAGTGCCCTTTTCCTCAGCACTGAGG 240
QY 1265 GAGGGGGCTTCTTTCCCTTCCCTCCCGGCGACAAGCTCCAGGGCAGGGGTGTCCCTCTGG 1324
Db 241 GAGGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGGCAGGGGTGTCCCTCTGG 300
QY 1325 GCGGCCAGCACTTCCCTCAGACACAACCTTCTCTGCTGCTCCAGTCGTGGGATCATCA 1384
Db 301 GCGGCCAGCACTTCCCTCAGACACAACCTTCTCTGCTGCTCCAGTCGTGGGATCATCA 360
QY 1385 CTTACCCACCCCGGCAAGTTCAAGACCAAAATCTCCAGCTGCCCCCTTCTGTTCCTCTGT 1444
Db 361 CTTACCCACCCCGGCAAGTTCAAGACCAAAATCTCCAGCTGCCCCCTTCTGTTCCTCTGT 420
QY 445 GTTGTCTGTAGCTGGGATGTCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 1504
Db 421 GTTGTCTGTAGCTGGGATGTCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 480
QY 1505 CTGACCCCTTCTTAATTCCTTAAGTCTAAAGATGATGAATTC 1546
Db 481 CTGACCCCTTCTTAATTCCTTAAGTCTAAAGATGATGAATTC 522

```

Search completed: January 22, 2003, 23:26:32
 Job time : 4855.54 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:31:07 ; Search time 354.534 Seconds
(without alignments)
9820.189 Million cell updates/sec

Title: US-09-888-358-2
Perfect score: 1546
Sequence: 1 ggcctaggcgctgcgagc.....gtctaaagatgatgaattc 1546

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Tc: number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002 : *

1:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1546	100.0	1546	AA171020	Human mitochondria
2	1500.4	97.1	1560	AA171020	Human uncoupling p
3	1421.4	91.9	1662	AA171019	Human ORFX ORF2547
4	1054.4	68.2	1114	AA171019	Human mitochondria
5	522	33.8	537	AA171019	Human immunogenic
6	522	33.8	537	AA171019	Human prostate cdn
7	522	33.8	537	AA171019	Human prostate-spe
8	522	33.8	537	AA171019	Human prostate-spe
9	522	33.8	537	AA171019	Prostate tumour an

10	522	33.8	537	24	ABL95186	Human CGI-69 cDNA
11	517.2	33.5	2170	24	ABK34735	Human cDNA for nov
12	470.4	30.4	704	24	ABK34735	Mouse ischaemic CO
13	445.4	28.8	452	24	ABK34735	Gene #1133 used to
14	437.4	28.3	758	24	ABL01419	Murine apoptosis r
15	242.2	15.7	6712	21	AA171019	Human ORFX ORF2547
16	218.4	14.1	240	24	ABK34735	Gene #1179 used to
17	185.8	12.0	452	24	ABL01419	Human ovarian can
18	172	11.1	275	24	ABL01419	Human reproductive
19	172	11.1	275	24	ABL01419	Human cDNA encodin
20	172	11.1	275	24	ABL01419	cDNA encoding nove
21	170.2	11.0	512	24	ABK34735	Rat sequence diffe
22	166	10.7	174	16	AA171019	Human gene signatu
23	151.4	9.8	1258	24	ABL01419	Human polynucleoti
24	142.6	9.2	397	21	AAH30252	Human colon cancer
25	128.8	8.3	1449	23	ABL04893	Drosophila melanog
26	128.8	8.3	3928	23	ABL04893	Drosophila melanog
27	122	7.9	422	22	AAK33810	Murine transport a
28	111.8	7.2	1445	23	ABL24583	Drosophila melanog
29	79.6	5.1	3645	23	ABL24583	Drosophila melanog
30	78.2	5.1	1529	23	ABL24583	Drosophila melanog
31	77.4	5.0	2516	24	ABL90270	Human polynucleoti
32	74.8	4.8	452	22	ABA59262	Human foetal liver
33	74.8	4.8	452	22	AAK07474	Human brain expres
34	74.8	4.8	452	22	AAK33259	Human bone marrow
35	74.8	4.8	452	22	AAI39053	Probe #7739 used t
36	74.8	4.8	452	24	ABS08091	Human genome-deriv
37	74.8	4.8	496	22	ABA59322	Human foetal liver
38	74.8	4.8	496	22	AAK07540	Human brain expres
39	74.8	4.8	496	22	AAK33336	Human bone marrow
40	74.8	4.8	496	22	AAI39123	Probe #7809 used t
41	74.8	4.8	496	24	ABS08171	Human genome-deriv
42	63.8	4.1	367	24	ABL01420	Murine apoptosis r
43	61	3.9	997	21	AAZ61628	cDNA encoding muri
44	61	3.9	997	21	AAZ61628	Skin cell cDNA, SE
45	61	3.9	997	24	ABL34713	Murine cDNA isolat

ALIGNMENTS

RESULT 1	
AA171020	
ID	AA171020 standard; cDNA; 1546 BP.
AC	AA171020;
XX	
DT	18-MAR-2002 (first entry)
DE	Human mitochondrial carrier protein CGI-69 cDNA.
XX	
KW	CGI-69; mitochondrial carrier protein; human; metabolic disease;
KW	obesity; cachexia; tumour; cancer; infection; immunomodulator;
KW	antitumour; viricide; antibacterial; anorectic; antidiabetic;
KW	brown adipose tissue; diagnosis; gene therapy; ss.
OS	Homo sapiens.
XX	
PH	Key
FT	CDS
FT	118..1173
XX	/*tag= a
XX	
XX	Location/Qualifiers
XX	27-DEC-2001.
XX	22-JUN-2001; 2001WO-US20117.
XX	22-JUN-2000; 2000US-213307P.
XX	(GETH) GENENTECH INC.
XX	(CURA-) CURAGEN CORP.

PI Lewin D, Adams SH, Yu XX;
 XX MPI; 2002-114569/15.
 XX
 XX CGI-69 polypeptides and polynucleotides useful for treating metabolic
 XX disorders including cachexia, obesity, diabetes and cancers -
 XX
 XX Disclosure; Page 11-12; 107pp; English.
 XX
 XX The present sequence is that of cDNA encoding human CGI-69 (see
 XX AM50570). Various CGI-69 clones were isolated from human liver
 XX upon PCR amplification and cloning. Some diverged from the
 XX previously known CGI-69 sequence in that they encoded an 8-amino
 XX acid insert preceded by a weak change, and were termed the "long
 XX version" isoform or CGI-69L (see AAF1019). The invention relates
 XX to a novel characterization of CGI-69 as a mitochondrial carrier
 XX protein, the discovery of CGI-69L, and the discovery of the
 XX alteration of the mitochondrial membrane potential through
 XX overexpression of carboxyl-RuAG-tagged CGI-69. CGI-69 (including
 XX splice variant) nucleic acids and proteins are useful for
 XX diagnosing and treating metabolic diseases in humans, including
 XX obesity, cachexia and diabetes. Evidence for an important function
 XX for CGI-69 in modifying mitochondrial membrane potential in brown
 XX adipose tissue (BAT) is presented. The mouse orthologue of CGI-69
 XX is up-regulated in cold-treated BAT. CGI-69 may be involved in
 XX cellular thermogenic uncoupling and, therefore, may be used to
 XX diagnose and treat specific perturbations in metabolic pathways.
 XX Altering the expression of CGI-69 through gene therapy provides a
 XX means of treating metabolic diseases, such as obesity or cachexia,
 XX or of increasing or decreasing body weight. Decreasing the
 XX activity of CGI-69 can be used to treat cachexia, tumours, cancers,
 XX viral infections and bacterial infections, while increasing its
 XX activity can be used to treat obesity, tumours, cancers, viral
 XX infections and bacterial infections (all claimed). Methods are
 XX also claimed for determining whether a compound up-regulates or
 XX down-regulates expression of a CGI-69 gene, and of screening for
 XX a mutation in the CGI-69 gene.
 XX
 XX Sequence 1546 BP; 272 A; 489 C; 459 G; 326 T; 0 other;

Db 421 GCCTTGGTGAAGATCGTGGACGAGGCGACAGGACCCCTCTGGAGCGGCTCCCGCC 480
 Qy 481 ACCCTGGTGATGACTGCGCCAGCTACCGCATCTACTTCACTGCTATGACCACTGAAG 540
 Db 481 ACCCTGGTGATGACTGCGCCAGCTACCGCATCTACTTCACTGCTATGACCACTGAAG 540
 Qy 541 GCCTTCTCTGTGTGGTGGAGCCCTGACCTCTGACCTTACGCCACCCATGTTGGCTGGCG 600
 Db 541 GCCTTCTCTGTGTGGTGGAGCCCTGACCTCTGACCTTACGCCACCCATGTTGGCTGGCG 600
 Qy 601 CTGGCCCGCTGGGACCCGCTGATGATGACCCCTCTGGAGCTTATGGGACAAAGCTG 660
 Db 601 CTGGCCCGCTGGGACCCGCTGATGATGACCCCTCTGGAGCTTATGGGACAAAGCTG 660
 Qy 661 CAGGCTCAGCATGTGTGCTACCGGAGCTGGTGTCTGTTGCAACTGCACTGCTGCTCAG 720
 Db 661 CAGGCTCAGCATGTGTGCTACCGGAGCTGGTGTCTGTTGCAACTGCACTGCTGCTCAG 720
 Qy 721 GGTGGCTGGGCTCACTGTGTCTGGGCTGGGCGCCCACTGCTTCGAGATGTCCTTC 780
 Db 721 GGTGGCTGGGCTCACTGTGTCTGGGCTGGGCGCCCACTGCTTCGAGATGTCCTTC 780
 Qy 781 TCAGCCCTGTACTGTTCAACTATGAGCTGTGAGAGCTGGTCAATGGGTTGAGGCG 840
 Db 781 TCAGCCCTGTACTGTTCAACTATGAGCTGTGAGAGCTGGTCAATGGGTTGAGGCG 840
 Qy 841 AAGGACCAAGATCTGTGGGCTGATGCTTTTGGCTGGTGGCATCTCAGGACCGTGGCT 900
 Db 841 AAGGACCAAGATCTGTGGGCTGATGCTTTTGGCTGGTGGCATCTCAGGACCGTGGCT 900
 Qy 901 GAGTGGCTGATCTACCTTCACTGCTGATGATGATGATGATGATGATGATGATGATG 960
 Db 901 GAGTGGCTGATCTACCTTCACTGCTGATGATGATGATGATGATGATGATGATG 960
 Qy 961 ATGGAGGCTGTGAGGTGAACCCCTGATGATGATGATGATGATGATGATGATGATG 1020
 Db 961 ATGGAGGCTGTGAGGTGAACCCCTGATGATGATGATGATGATGATGATGATGATG 1020
 Qy 1021 ATCCGGGC TACTGGGACCAAGGACTCTTTCAGGCTTCTTCAGGCTTCTTCAGATCAT 1080
 Db 1021 ATCCGGGC TACTGGGACCAAGGACTCTTTCAGGCTTCTTCAGGCTTCTTCAGATCAT 1080
 Qy 1081 GTTGCCCTCTCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 Db 1081 GTTGCCCTCTCTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 Qy 1141 AGGCTGAACAGGACCGCTTCTGGGCGGCTGAAAGGGGCAAGGAGCAAGACCCGCTC 1200
 Db 1141 AGGCTGAACAGGACCGCTTCTGGGCGGCTGAAAGGGGCAAGGAGCAAGACCCGCTC 1200
 Qy 1201 TCTCCACCGATGGGAGAGGCGAGGAGAGACCCAGCAAGTGCCTTCTTCAGCACT 1260
 Db 1201 TCTCCACCGATGGGAGAGGCGAGGAGAGACCCAGCAAGTGCCTTCTTCAGCACT 1260
 Qy 1261 GAGGAGGAGGCTTGTTCCTTCCCTCCGCGGCAAGCTCCAGGCGAGGCTGTCCCT 1320
 Db 1261 GAGGAGGAGGCTTGTTCCTTCCCTCCGCGGCAAGCTCCAGGCGAGGCTGTCCCT 1320
 Qy 1321 CTGGGCGCCAGCACTTCTTCAGACAACTTCTTCGCTGCTCCAGTGTGGGATC 1380
 Db 1321 CTGGGCGCCAGCACTTCTTCAGACAACTTCTTCGCTGCTCCAGTGTGGGATC 1380
 Qy 1381 ATCACTTACCCACCCCAAGTTCAGAGCAAAATCTTCCAGCTGCCCCCTTCGTGTTC 1440
 Db 1381 ATCACTTACCCACCCCAAGTTCAGAGCAAAATCTTCCAGCTGCCCCCTTCGTGTTC 1440
 Qy 1441 CTGTGTTGCTGATGCTGGGCTGCTTCAGAGCAAAAGCCCTCAGCTGTGTGAT 1500
 Db 1441 CTGTGTTGCTGATGCTGGGCTGCTTCAGAGCAAAAGCCCTCAGCTGTGTGAT 1500
 Qy 1501 CTCCCTGACCCCTGTTAAATTCCTTAAGTCTAAAGATGATGATGATGATGATGATG 1546

Query Match 100.0%; Score 1546; DB 24; Length 1546;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTAGTGGCTGCGGCGCGGAGCCAGGCGGAGCGAGCTAAATGGGCGCGCT 60
 Db 1 GGCTAGTGGCTGCGGCGCGGAGCCAGGCGGAGCGAGCTAAATGGGCGCGCT 60
 Qy 61 GGCCCTGGGCGCGGCGCGGAGCCAGGCGGAGCGAGCTAAATGGGCGCGCT 120
 Db 61 GGCCCTGGGCGCGGCGCGGAGCCAGGCGGAGCGAGCTAAATGGGCGCGCT 120
 Qy 121 GCTGACACGAGCCCTGGGCGGATCAGCCCTCCAGCAATGCTGAGGCGCGCG 180
 Db 121 GCTGACACGAGCCCTGGGCGGATCAGCCCTCCAGCAATGCTGAGGCGCGCG 180
 Qy 181 GCTGTGTTACCTCTCTTTCATGACACCCCTGAGCTGAGGCTGAGGCTGAGGCT 240
 Db 181 GCTGTGTTACCTCTCTTTCATGACACCCCTGAGCTGAGGCTGAGGCTGAGGCT 240
 Qy 241 CAGGCGCCCTCCATGGCAGGAGCTGATGCTTCTTCAGACTGTGGAGCTTCTCTAT 300
 Db 241 CAGGCGCCCTCCATGGCAGGAGCTGATGCTTCTTCAGACTGTGGAGCTTCTCTAT 300
 Qy 301 ACCAAATGGAAGTGGCTCTGATGCAATGCTGCTGGAGCTCTGTAACCTGTGCCA 360
 Db 301 ACCAAATGGAAGTGGCTCTGATGCAATGCTGCTGGAGCTCTGTAACCTGTGCCA 360
 Qy 361 AATGGTGGCCGCTGTGCCACCTGGTTTCAAGACCCCTACCCGCTTCACTGGCACTGAT 420
 Db 361 AATGGTGGCCGCTGTGCCACCTGGTTTCAAGACCCCTACCCGCTTCACTGGCACTGAT 420
 Qy 421 GCCTTGTGAAGATCGTGGGCGAGGCGACGAGGACCCCTCTGGAGCGGCTCCCGCC 480

Db 1501 CTCCCTGACCCCTTGTTAATTCCTTAAGCTAAAGATGATGAACCTTC 1546

RESULT 2

AAC90452

ID AAC90452 standard; cDNA; 1560 bp.

AAC90452:

DT 12-MAR-2001 (first entry)

Human uncoupling protein cDNA #1.

Human; uncoupling protein; immunosuppressive; antiarthritic;
anti-rheumatic; antiproliferative; cardiatic; vasotropic;
cerebroprotective; neuroprotective, antibacterial; ophthalmological;
gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;
gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
infertility; ss.

XX
:5C
omo sapiens.

XX
PN
WO200061614-A2

19-OCT-2000.

06-APR-2000;

09-APR-1999; 99US-0128701

99US-0142821.
99US-0149448.

12-NOV-1999; 99US-0164751.

PA (HUMTA-) HUMAN GENOME SCI INC.

RU BEN SM, NI J, KOMATSUOLIS

PK
R
NPI; 2000-656322/63.
P-PSDB; AAB50378.

Uncoupling proteins and nucleic acid sequences encoding

immune system, cardiovascular and gastrointestinal diseases, preventing and treating proliferative, neoplastic diseases, for detecting, preventing and treating proliferative, neoplastic diseases.

Claim 1; Page 303; 343pp; English.

The present sequence is one of eighteen isolated nucleotides

encoding uncoupling proteins. The nucleotide sequences may be used for the detection of various disorders such as cancer, for chromosome identification, as chromosome markers and for numerous other diagnostic or research purposes. The uncoupling protein encoded by the nucleotide sequences may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders, wounds, infectious diseases, thrombosis, arthritis, and infertility.

Sequence 1560 BP; 286 A; 492 C; 459 G; 323 T; 0 other:

```
Query Match          97.1%; Score 1500.4; DB 21; Length 1560;
Best Local Similarity 98.8%; Pred. NO. 0;
Matches 1510; Conservative 0; Mismatches 18; Indels 0;
```

14 GCGAGCGCGGAGGCCACGAGGGCGGACGGA CGTATTCGCCCCCCCCCCCCCCCCCCCC

16 GCGGCGGGACCGGCACAGGGCGGAGCCGGTATGGCCCCCGCTGCCCTGCCCCC

74 GCGCCGACGAGCACCAGCCTAGGCCAGGA CTGAGGCTTCAGATGGCTCA CCGTCTCT

76 GGGCCGACGAGCACCGCTTAGAGCCAGGACTGAAGCTTCACATGCGCTTACCGCGTCC

134 CTGGGGCATCAGCCCCCTCCAGCAATGTTGGCTCAGGCACCGGGATTGCCAACA

136 CTGGGGCATCAGCCCCCTCCAGCAATGTTGGCTTCAGGCACCGGGTGTGCTTGCTT

QY	194	CTCTCTTCATGACACCCCTGGACGTGTGTGAAGGTTTGCCTGCGAGTCTCAGCGGCCCTTCCA	255
Db	196	CTCTCTTCATGACACCCCTGGACGTGTGTGAAGGTTTGCCTGCGAGTCTCAGCGGCCCTTCCA	255
QY	254	TGGCCAGGAGCTGATGCTCTTCAGACTCTGTGAGCCCTCTCTTATACCAAAATGGAAGT	313
Db	256	TGGCCAGGAGCTGATGCTCTTCAGACTCTGTGAGCCCTCTCTTATACCAAAATGGAAGT	315
QY	314	GCCTCTCTGTATTTGCAATGCTCTCTGGAGCCTCTGTACTGTGACCCATATGGTGTCCCGCT	373
Db	316	GCCTCTCTGTATTTGCAATGCTCTCTGGAGCCTCTGTACTGTGACCCATATGGTGTCCCGCT	375
QY	374	GTGCCACCTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATGATGSCCTTGTGGAACA	433
Db	376	GTGCCACCTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATGATGSCCTTGTGGAAGA	435
QY	434	TCGTGAGGCACAGAGGCACAGACCCCTCTGGAGGCTCTCCCGCCACCCCTGTGTATGA	493
Db	436	TCGTGAGGCACAGAGGCACAGACCCCTCTGGAGGCTCTCCCGCCACCCCTGTGTATGA	495
QY	494	CTGTGCACGACTACCGCCATCTACTGCTTATGACCAACTGAAGGCTTCTCTGTGT	553
Db	496	CTGTGCACGACTACCGCCATCTACTGCTTATGACCAACTGAAGGCTTCTCTGTGT	555
QY	554	GTGAGCCCTGACCTCTGACCTCTACGACACTGTGGCTGTGGCGCTGCGCCCGCTGG	613
Db	556	GTGAGCCCTGACCTCTGACCTCTACGACACCCATGTGGCTGTGGCGCTGCGCCCGCTGG	615
QY	614	GCAACCGTGACTGTGATCAGGCCCTGTGAGCTTTATGGGACAAGCTCAGGCTCAGCATG	673
Db	616	GCAACCGTGACTGTGATCAGGCCCTGTGAGCTTTATGGGACAAGCTCAGGCTCAGCATG	675
QY	674	TGTGCTTACCGGAGCTCGGTGCTGTGTGTGAACCTGACAGTGTGCTCAGGCTGTGCGCT	733
Db	676	TGTGCTTACCGGAGCTCGGTGCTGTGTGTGAACCTGACAGTGTGCTCAGGCTGTGCGCT	735
QY	734	CATGTGCTGTGGGTGTGGGGCCCACTGCGCTTCTGAGATGTGCCCTCTCAGAGCCCTGTACT	793
Db	736	CATGTGCTGTGGGTGTGGGGCCCACTGCGCTTCTGAGATGTGCCCTCTCAGAGCCCTGTACT	795
QY	794	GGTTCAACTATGAGCTGTGTGAAGAGCTGGCTCAATGGGTTCAGCGCCGAAGGACCAAGACTT	853
Db	796	GGTTCAACTATGAGCTGTGTGAAGAGCTGGCTCAATGGGTTCAGCGCCGAAGGACCAAGACTT	855
QY	854	CTGTGGGCATCAGCTTTTGTGCTGTGGTCACTCTGAGAGAGGTGGCTCAGTGTGTGATCTC	913
Db	856	CTGTGGGCATCAGCTTTTGTGCTGTGGTCACTCTGAGAGAGGTGGCTCAGTGTGTGATCTC	915
QY	914	TACCTTTGACGTGTGAAGACCAACGCGAGCTGGCTCTGGAGCGATGAGAGGCTGTGA	973
Db	916	TACCTTTGACGTGTGTGAAGACCAACGCGAGCTGGCTCTGGAGCGATGAGAGGCTGTGA	975
QY	974	GAGTGACACCCCTGCACTGTGACTTCACTGTGCTGTGCTGTGCGAGGATTCGGGCGGAGT	1033
Db	976	GAGTGACACCCCTGCACTGTGACTTCACTGTGCTGTGCTGTGCGAGGATTCGGGCGGAGT	1035
QY	1034	CGGGCACCAAGGAGCTCTTTTTCAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCTTCCT	1093
Db	1036	CGGGCACCAAGGAGCTCTTTTTCAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCTTCCT	1095
QY	1094	GTGCCATCATGATCAGCACCTATGAGTTTCGGCAAAAGCTTCTTTCAGAGGCTGTGAACAGG	1153
Db	1096	GTGCCATCATGATCAGCACCTATGAGTTTCGGCAAAAGCTTCTTTCAGAGGCTGTGAACAGG	1155
QY	1154	ACCGGCTTCTGGGCGGCTGAAAGGGGCAGGAGCAAGGACCCCTCTCTCCACGGATG	1213
Db	1156	ACCGGCTTCTGGGCGGCTGAAAGGGGCAGGAGCAAGGACCCCTCTCTCCACGGATG	1215
QY	1214	GGGAGAGGGCAGGAGGAGACCCAGCCAGTGTGCTTTTCTCAGACTGAAGGAGGGGCT	1273
Db	1216	GGGAGAGGGCAGGAGGAGACCCAGCCAGTGTGCTTTTCTCAGACTGAAGGAGGGGCT	1275

QY 1274 TGTTTCCTTCCCTCCCGGCGAAGCTCCAGGCGAGGGCTGTCCCTCTGGGCGGCCAG 1333
 Db 1276 TGTTTCCTTCCCTCCCGGCGAAGCTCCAGGCGAGGGCTGTCCCTCTGGGCGGCCAG 1335
 QY 1334 CACTTCTCAGACAACTTCTCTGCTGCTCAGTCTGCTGGGATCATCACTTACCAC 1393
 Db 1336 CACTTCTCAGACAACTTCTCTGCTGCTCAGTCTGCTGGGATCATCACTTACCAC 1395
 QY 1394 CCCCAGTTCAGACCAAACTTCCAGCTGCCCTCGTGTTCCTCTGTTGTTGTT 1453
 Db 1396 CCCCAGTTCAGACCAAACTTCCAGCTGCCCTCGTGTTCCTCTGTTGTTGTT 1455
 QY 1454 AGCTGGGATGTTCCAGAACCAAGAGCCCTCAGCTGGTGTAGTCTCCCTGACCTT 1513
 Db 1456 AGCTGGGATGTTCCAGAACCAAGAGCCCTCAGCTGGTGTAGTCTCCCTGACCTT 1515
 QY 1514 GTTAATTCCTTAAGTCTTAAGATGATGAA 1542
 Db 1516 GTTAATTCCTTAAGTCTTAAGATGATGAA 1544

RESULT 3
 AAC76992
 ID AAC76992 standard; cDNA; 1662 BP.
 AC AAC76992;
 XX
 DT 08-FEB-2001 (first entry)
 DE
 XX
 DE Human ORF2547 polynucleotide sequence SEQ ID NO:5093.
 XX
 DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnerable; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Shimketa RA. Leach M;
 PI WPI; 2000-602362/57.
 DR P-PSDB; AAB42783.
 DR
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5; Page 4281-4282; 5507pp; English.
 XX
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; neuroprotective;
 CC antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy. The
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis, and as a contraceptive.
 XX

Sequence 1662 BP; 336 A; 516 C; 468 G; 341 T; 1 other;

Query Match 91.9%; Score 1421.4; DB 21; Length 1662;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1530; Conservative 0; Mismatches 16; Indels 59; Gaps 4;

QY 1 GGCTAGGTGCGTGGAGCGCGG-----CAGCCAGAGGCGGAGCGAGCTAATGGGCC 56
 Db 2 GGCTAGGTGCGTGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 61
 QY 57 GCGTGGCCCTGGGCG 103
 Db 62 GCGTGGCCCTGGGCG 121
 QY 104 ACTGAGCTTCAAGATGGCTGACAGGACCGCTGGCGGCATCAGCCCGCTCCAGCAATGG 163
 Db 122 ACTGAGCTTCAAGATGGCTGACAGGACCGCTGGCGGCATCAGCCCGCTCCAGCAATGG 181
 QY 164 TGGCTCTCAGGACCG 223
 Db 182 TGGCTCTCAGGACCG 241
 QY 224 AGGTTGCGCTGCAATCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 283
 Db 242 AGGTTGCGCTGCAATCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
 QY 284 TGTGGAG-----CCTCTCTATACCAATGAAGTGCCTCC 319
 Db 302 TGTGGAGCGCTCTCTATACCAATGAAGTGCCTCCCTCTCTATACCAATGAAGTGCCTCC 361
 QY 320 TGTATTTGCAATGCTGCTGGAGCGCTGTACCTGTGCGCCCAATGGTGCCTGTGCGCA 379
 Db 362 TGTATTTGCAATGCTGCTGGAGCGCTGTACCTGTGCGCCCAATGGTGCCTGTGCGCA 421
 QY 380 CCGTGGTTCAAGACCGCTACCGCTTCACTGGCACCATGATGCTTCTGTGAAGATCTGTA 439
 Db 422 CCGTGGTTCAAGACCGCTACCGCTTCACTGGCACCATGATGCTTCTGTGAAGATCTGTA 481
 QY 440 GGCACGAGGGCACAGGACCGCTTGGAGCGGCGTCCCGCCACCGCTGGTGTGCTGCG 499
 Db 482 GGCACGAGGGCACAGGACCGCTTGGAGCGGCGTCCCGCCACCGCTGGTGTGCTGCG 541
 QY 500 CAGCTACCGCGCATCTACTTCACTGCTATGACCACTGAGGCGCTTCTGTGTGCTGCG 559
 Db 542 CAGCTACCGCGCATCTACTTCACTGCTATGACCACTGAGGCGCTTCTGTGTGCTGCG 601
 QY 560 CCGTGAACCTTGACCTTACGACACCGCATGCTGGCGCGCTGGCGCGCGCTGGCGCACCG 619
 Db 602 CCGTGAACCTTGACCTTACGACACCGCATGCTGGCGCGCTGGCGCGCGCTGGCGCACCG 661
 QY 620 TGACTGTGATGAGCGCGCTGAGCTTATGGGCAAAAGCTGAGGCTCAGCATGTGCTGT 679

Db 662 TGACTGTGATCAGCCCTGGAGCTTATGCCGACAAAGCTGCAGGCTCAGCATGTGTGT 721
 QY 680 ACCGGAGCTGGGTGGCTGTGTTCGAATGCAAGTGGCTCAGGCTGGCTGGCTCACTGT 739
 Db 722 ACCGGAGCTGGGTGGCTGTGTTCGAATGCAAGTGGCTCAGGCTGGCTGGCTCACTGT 781
 QY 740 GGCTGGCTGGGCCCCCTGCTTCCAGATGTGCCCTTCTC-----782
 Db 782 GGCTGGCTGGGCCCCCTGCTTCCAGATGTGCCCTTCTCAGTGCATCCCCACCCC 841
 QY 783 -AGCCCTGTACTGTTTCAACTATGAGCTGTGAAGAGCTGGCTCAATGGTTCAGGCGGA 841
 Db 842 AAGCCCTGTACTGTTTCAACTATGAGCTGTGAAGAGCTGGCTCAATGGTTCAGGCGGA 901
 QY 842 AGGACAGACTTCTGTGGGATGAGCTTGTGGCTGGTGGCTCAGGACGGTGGCTG 901
 Db 902 AGGACAGACTTCTGTGGGATGAGCTTGTGGCTGGTGGCTCAGGACGGTGGCTG 961
 QY 902 CAGTGTGACTTACCTTTTACCTGTTAAGACCCCAAGCCCAAGCCAGTCTCTGGAGCGA 961
 Db 962 CAGTGTGACTTACCTTTTACCTGTTAAGACCCCAAGCCCAAGCCAGTCTCTGGAGCGA 1021
 QY 962 TGAGAGCTGTGAGAGTGAACCCCTCGCATGTGACTTCCACTGGCTGGCTGGAGCGA 1021
 Db 1022 TGAGAGCTGTGAGAGTGAACCCCTCGCATGTGACTTCCACTGGCTGGCTGGAGCGA 1081
 QY 1022 TCCGGGCGAGTGGGGACCAAGGACTTTTTCAGAGTCTTCTTCGGATCATCAAG 1081
 Db 1082 TCCGGGCGAGTGGGGACCAAGGACTTTTTCAGAGTCTTCTTCGGATCATCAAG 1141
 QY 1082 CTGCCCCCTCTCTGCTCCATCATGATCAGCAGTATGAGTTCGGGAAAGCTTCTCCAGA 1141
 Db 1142 CTGCCCCCTCTCTGCTCCATCATGATCAGCAGTATGAGTTCGGGAAAGCTTCTCCAGA 1201
 QY 1142 GGCTGAACAGAGACCGGCTTCTGGGGGCTGAAAGGGCAAGAGGCAAGGACCCCTCT 1201
 Db 1202 GGCTGAACAGAGACCGGCTTCTGGGGGCTGAAAGGGCAAGAGGCAAGGACCCCTCT 1261
 QY 1202 CTCCACGAGTGGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261
 Db 1262 CTCCACGAGTGGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1321
 QY 1262 AGGAGGGGGCTTCTTCTCCCTCCCTCCGCGCAAGCTTCCAGTGGCTGGGATCA 1321
 Db 1322 AGGAGGGGGCTTCTTCTCCCTCCCTCCGCGCAAGCTTCCAGTGGCTGGGATCA 1381
 QY 1322 TGGGGGCGCCAGCAGCTTCTCAGACAACTTCTTCTGCTGCTCAGTGGGATCA 1381
 Db 1382 TGGGGGCGCCAGCAGCTTCTCAGACAACTTCTTCTGCTGCTCAGTGGGATCA 1441
 QY 1382 TCACCTACCCACCCCAAGTTCAGACCAATCTTCCAGTGGGATCA 1441
 Db 1442 TCACCTACCCACCCCAAGTTCAGACCAATCTTCCAGTGGGATCA 1501
 QY 1442 TGTGTTGTGTAGCTGGGATGCTTCCAGGACCAAGAGCCCTCAGCTGGTGTATC 1501
 Db 1502 TGTGTTGTGTAGCTGGGATGCTTCCAGGACCAAGAGCCCTCAGCTGGTGTATC 1561
 QY 1502 TCCCTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1546
 Db 1562 TCCCTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1606

RESULT 4

AAI71019

ID AAI71019 standard; cDNA; 1114 BP.

XX AAI71019;

AC AAI71019;

XX AAI71019;

DT 18-MAR-2002 (first entry)

DE Human mitochondrial carrier protein CGI-69L cDNA.

XX

KW CGI-69L; mitochondrial carrier protein; human; metabolic disease;
 KW obesity; cachexia; tumor; cancer; infection; immunomodulator;
 KW antitumor; viricide; antibacterial; anorectic; anti-diabetic;
 KW brown adipose tissue; diagnosis; gene therapy; splice variant; ss.
 XX Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 14..1093
 FT /*tag= a
 XX
 XX WO200198355-A2.
 PD 27-DEC-2001.
 XX
 PF 22-JUN-2001; 2001WO-US20117.
 XX
 PR 22-JUN-2000; 2000US-213307P.
 XX
 PA (GETH) GENENTECH INC.
 PA (CURA-) CURAGEN CORP.
 XX
 PI Lewin D, Adams SH, Yu XX;
 XX WPI; 2002-114569/15.
 DR P-PSDB; AAM50569.
 XX

CGI-69 polypeptides and polynucleotides useful for treating metabolic disorders including cachexia, obesity, diabetes and cancers -
 Claim 1(a); Page 9-10; 107pp; English.

The present sequence is that of cDNA encoding human CGI-69L (see AAM50569), a novel splice variant of human CGI-69. Various CGI-69 clones were isolated from human liver upon PCR amplification and cloning. Numerous clones diverged from the previously known CGI-69 sequence in that they encoded an 8-amino acid insert preceded by a W44L change, and were termed the 'long version' isoform or CGI-69L. The invention relates to a novel characterization of CGI-69 as a mitochondrial carrier protein, the discovery of CGI-69L as a discovery of the alteration of the mitochondrial membrane potential through overexpression of carboxy-FLAG-tagged CGI-69. CGI-69 (including splice variant) nucleic acids and proteins are useful for diagnosing and treating metabolic diseases in humans, including obesity, cachexia and diabetes. Evidence for an important function for CGI-69 in modifying mitochondrial membrane potential in brown adipose tissue (BAT) is presented. The mouse orthologue of CGI-69 is up-regulated in cold-treated BAT. CGI-69 may be involved in cellular thermogenic uncoupling and, therefore, may be used to diagnose and treat specific perturbations in metabolic pathways. Altering the expression of CGI-69 through gene therapy provides a means of treating metabolic diseases, such as obesity or cachexia, or of increasing or decreasing body weight. Decreasing the activity of CGI-69 can be used to treat cachexia, tumors, cancers, viral infections and bacterial infections (all claimed). Methods are also claimed for determining whether a compound up-regulates or down-regulates expression of a CGI-69 gene, and of screening for a mutation in the CGI-69 gene.

Sequence 1114 BP; 194 A; 349 C; 332 G; 239 T; 0 other;

Query Match 68.2%; Score 1054.4; DB 24; Length 1114;

Best Local Similarity 97.8%; Pred. No. 8.9e-250;

Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 105 CTGAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGGT 164
 Db 1 CTGAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGGT 60

QY 165 GGCTCAGGACCGGGCTGTGGTACCTCTCTTCATGACCCCTGGACCTGGTGA 224

Db	61	GGCCTCAGGCACCGGGGCTGTGGTTACTCTCTCTTCTATGACACACCCCTGGACGTGGTGAA	120
Qy	225	GGTTTGGCTCGAGTCTCAGGGCCCTCCATGGCCAGCAGCTGATGCTTCCTCCACGACT	284
Db	121	GGTTTGGCTCGAGTCTCAGGGCCCTCCATGGCCAGCAGCTGATGCTTCCTCCACGACT	180
Qy	285	GTGGAGCCTCTCTCTATACCAAT-----GGAGTGGCTCTCT	320
Db	181	GTGGAGCCTCTCTATACCAATTTGCCCTCTCTCTCCAAATCCACAGGGAAGTGCCTCT	240
Qy	321	GTATTGCAATGGTCTCTGGAGCCTCTGACTGTGCCCAATGGTGGCCCGCTGTGCCAC	380
Db	241	GTATTGCAATGGTCTCTGGAGCCTCTGACTGTGCCCAATGGTGGCCCGCTGTGCCAC	300
Qy	381	CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCCITTCGTGAAGATCGTGAG	440
Db	301	CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCCITTCGTGAAGATCGTGAG	360
Qy	441	GCAGAGGACACAGAGCCCTCTGGAGCGCCTCCCGCGACCTCTGTGATGACTGTGCC	500
Db	361	GCAGAGGACACAGAGCCCTCTGGAGCGCCTCCCGCGACCTCTGTGATGACTGTGCC	420
Qy	501	AGCTACCGGCATCTACTTCACTGCGCTATGACCAACTGAAGCGCTTCTCTGTGGTGCAGC	560
Db	421	AGCTACCGGCATCTACTTCACTGCGCTATGACCAACTGAAGCGCTTCTCTGTGGTGCAGC	480
Qy	561	CCTGACTCTGACTCTTACGCACCCATGFGGTGGCGCTGGCCCGCTGGGCACCGT	620
Db	481	CCTGACTCTGACTCTTACGCACCCATGFGGTGGCGCTGGCCCGCTGGGCACCGT	540
Qy	621	GACTGTGATCAGCCCCCTGGAGCTTATGGCGCAAGCTGCAGGCTCAGCATGTGTCTGA	680
Db	541	GACTGTGATCAGCCCCCTGGAGCTTATGGCGCAAGCTGCAGGCTCAGCATGTGTCTGA	600
Qy	681	CCGGAGCTGGTGCTGTGTGTGAACTGCAGTGGCTCAGGCTGGCTGGCGCTACTCTGTG	740
Db	601	CCGGAGCTGGTGCTGTGTGTGAACTGCAGTGGCTCAGGCTGGCTGGCGCTACTCTGTG	660
Qy	741	GCTGGGCTGGGGCCCACTGCCCTTCAGATGTGCCCTCTCAGCCCTGTACTGTCTCAA	800
Db	661	GCTGGGCTGGGGCCCACTGCCCTTCAGATGTGCCCTCTCAGCCCTGTACTGTCTCAA	720
Qy	801	CTATGAGCTGGTGAAGAGCTGGTCAATGGGTTCAAGCCGGAAGACACAGACTTCTGTGGG	860
Db	721	CTATGAGCTGGTGAAGAGCTGGTCAATGGGTTCAAGCCGGAAGACACAGACTTCTGTGGG	780
Qy	861	CATGAGCTTTGTGCTGGTGGCATCTCAGGAGCGTGGCTGCAGTGTCTACTCTACCTTT	920
Db	781	CATGAGCTTTGTGCTGGTGGCATCTCAGGAGCGTGGCTGCAGTGTCTACTCTACCTTT	840
Qy	921	TGACGTGTGTAAGACCCCAAGCCAGTTCCTCTGGAGCGATGGAGCTGTGAGAGTGAA	980
Db	841	TGACGTGTGTAAGACCCCAAGCCAGTTCCTCTGGAGCGATGGAGCTGTGAGAGTGAA	900
Qy	981	CCCCTCGATGTGACTTCACTGGCTGTCTGTGGAGGATCCGGCCCGAGTGGGGCAC	1040
Db	901	CCCCTCGATGTGACTTCACTGGCTGTCTGTGGAGGATCCGGCCCGAGTGGGGCAC	960
Qy	1041	CAAGGACTCTTTGCGAGCTTCTCTCTCGGATCATCAAGGCTGGCCCTCTCTGTGCCAT	1100
Db	961	CAAGGACTCTTTGCGAGCTTCTCTCTCGGATCATCAAGGCTGGCCCTCTCTGTGCCAT	1020
Qy	1101	CATGATCAGCACCTTATGAGTTGGCAAAAGCTTCTTCAGAGGCTCAACACAGGACCGGCT	1160
Db	1021	CATGATCAGCACCTTATGAGTTGGCAAAAGCTTCTTCAGAGGCTCAACACAGGACCGGCT	1080
Qy	1161	TTTGGGCGGTGAAGGGGCAAGAGGCAAGGAC	1194
Db	1081	TTTGGGCGGTGAAGGGGCAAGAGGCAAGGAC	1114

RESULT 5
AAA06606

AA06606 standard; cDNA; 537 BP.	XX
AA06606;	XX
13-JUN-2000 (first entry)	XX
Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:387.	XX
Human; prostate cancer; diagnosis; tumour; gene therapy; detection;	XX
immunogenic; cytostatic; vaccine; ss.	XX
Homo sapiens.	XX
W0200004149-A2.	XX
27-JAN-2000.	XX
14-JUL-1999; 99WO-US15838.	XX
14-JUL-1998; 98US-0115453.	XX
14-JUL-1998; 98US-0116134.	XX
23-SEP-1998; 98US-0159812.	XX
23-SEP-1998; 98US-0159822.	XX
15-JAN-1999; 99US-0232149.	XX
15-JAN-1999; 99US-0232880.	XX
09-APR-1999; 99US-0288946.	XX
{CORI-} CORIXA CORP.	XX
Dillon DC, Harlocker SL, Yudi J, Xu J, Mitcham JL;	XX
WPI; 2000-171269/15.	XX
New polypeptide useful for treating and diagnosing prostate cancer	XX
comprises an immunogenic portion of prostate tumor protein -	XX
Claim 50; Page 235-236; 263pp; English.	XX
The present invention describes isolated polypeptides, comprising an	XX
immunogenic portion of a prostate tumour protein (PTP). The polypeptides	XX
and polynucleotides encoding them have cytostatic activity and can be	XX
used in vaccines and in gene therapy. The polypeptides and	XX
polynucleotides encoding them, antigen presenting cells which express	XX
the polypeptides, antibodies against the polypeptides and v ccines	XX
comprising them can be used for inhibiting the development of prostate	XX
cancer in a patient. The polypeptides can be used to generate antibodies	XX
or anti-idiotypic antibodies for passive immuno therapy. A portion of	XX
the polynucleotides encoding the polypeptides can be used as a probe or	XX
to modulate the expression of the polypeptides. AAA06241 to AAA06691 an	XX
AAV82000 to AAV82020 represent sequences used in the exemplification of	XX
the present invention.	XX
Sequence 537 BP; 116 A; 167 C; 138 G; 116 T. 0 Other.	XX

RESULT 5
AAA06606

QY 1265 GAGGGGCTTGTTCCTTCCCTCCCGGCGAACAAGCTCCAGGCGAGGGCTGTCCCTCTGG 1324
DB 241 GAGGGGCTTGTTCCTTCCCTCCCGGCGAACAAGCTCCAGGCGAGGGCTGTCCCTCTGG 300
QY 1325 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATCATCA 1384
DB 301 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATCATCA 360
QY 1385 CTTACCCACCCCAAGTTCAAGACCAAACTTCCAGCTGCCCTTCTGTTTCCCTGT 1444
DB 361 CTTACCCACCCCAAGTTCAAGACCAAACTTCCAGCTGCCCTTCTGTTTCCCTGT 420
QY 1445 GTTTCCTGTAGCTGGGCATGCTCCAGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 1504
DB 421 GTTTCCTGTAGCTGGGCATGCTCCAGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 480
QY 1505 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1546
DB 181 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 522

RESULT 6
AAH93722
ID AAH93722 standard; cDNA; 537 BP.
XX
AC AAH93722;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human prostate cDNA sequence #357.
XX
KW Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS Homo sapiens.
XX
PW WO200173032-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US09919.
XX
PR 27-MAR-2000; 2000US-0536857.
XX
PR 09-MAY-2000; 2000US-0568100.
XX
PR 12-MAY-2000; 2000US-0570737.
XX
PR 13-JUN-2000; 2000US-0593793.
XX
PR 27-JUN-2000; 2000US-0605783.
XX
PF 10-AUG-2000; 2000US-0636215.
XX
P 19-AUG-2000; 2000US-0651236.
XX
PK 06-SEP-2000; 2000US-0657279.
XX
PR 02-OCT-2000; 2000US-0679426.
XX
PR 10-OCT-2000; 2000US-0685166.
XX
PA (CORI-) CORIXA CORP.
XX
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky IAW, Hepler WT, Henderson RA;
XX
DR WPI; 2001-639232/73.
XX
XX

PT New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS Claim 1; Page 361; 579pp; English.
XX
XX The invention relates to isolated prostate-specific
CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC antibodies raised against the polypeptides (or antigenic epitopes
CC derived from them) and antigen-presenting cells expressing the
CC polypeptides. The antibodies are useful for detecting the presence of
CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC the antigen-presenting cells are useful for stimulating and/or expanding

CC T cells specific for a tumour protein, and for inhibiting the development
CC of cancer especially prostate cancer. Compositions comprising the
CC polynucleotide and/or polypeptide are useful for stimulating an immune
CC response, and for treating cancer. The oligonucleotide is useful for
CC detecting cancer. The present sequence is a prostate specific
CC polynucleotide of the invention.

XX
SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 33.8%; Score 522; DB 22; Length 537;
Best Local Similarity 100.0%; Pred. No. 8.7e-119;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTGGGACACCAAGGACATCTTTCAGGCTTCTTCTCGATCATCAAGGCTG 1084

DB 1 GGGCCGAGTGGGACACCAAGGACATCTTTCAGGCTTCTTCTCGATCATCAAGGCTG 60

QY 1085 CCCCCTCTGTGCCATCATGATCAGCACTATGATTTGCGCAAAAGCTTCTTCAGAGGC 1144

DB 61 CCCCCTCTGTGCCATCATGATCAGCACTATGATTTGCGCAAAAGCTTCTTCAGAGGC 120

QY 1145 TGAACACAGGACCGGCTTCTTGGGCGCTGAAAAGGGCAAGAGGCAAGGCTTCTTC 1204

DB 121 TGAACACAGGACCGGCTTCTTGGGCGCTGAAAAGGGCAAGAGGCAAGGCTTCTTC 180

QY 1205 CCAGCGATGGGAGAGGCGAGGAGAGACCCAGCAAGTGCCTTCTCAGCACTGAGG 1264

DB 181 CCAGCGATGGGAGAGGCGAGGAGAGACCCAGCAAGTGCCTTCTCAGCACTGAGG 240

QY 1265 GAGGGGGCTTGTTCCTTCCCTCCCGCGACAAGCTCCAGGCGAGGCTGTCCCTCTGG 1324

DB 241 GAGGGGGCTTGTTCCTTCCCTCCCGCGACAAGCTCCAGGCGAGGCTGTCCCTCTGG 300

QY 1325 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATCATCA 1384

DB 301 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATCATCA 360

QY 1385 CTTACCCACCCCAAGTTCAAGACCAAACTTTCAGCTGCCCTTCTGTTTCCCTGT 1444

DB 361 CTTACCCACCCCAAGTTCAAGACCAAACTTTCAGCTGCCCTTCTGTTTCCCTGT 420

QY 1445 GTTTCCTGTAGCTGGGCATGCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 1504

DB 421 GTTTCCTGTAGCTGGGCATGCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 480

QY 1505 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1546

DB 481 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 522

RESULT 7

AAH93722

ID AAH93722 standard; cDNA; 537 BP.

XX

AC AAH93722;

XX

DT 04-OCT-2001 (first entry)

XX

DE Human prostate-specific cDNA sequence CGI-69.

XX

KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;

XX

KW cytostatic; gene therapy; metastasis; ss.

XX

OS Homo sapiens.

XX

PN WO200151633-A2.

XX

PD 19-JUL-2001.

XX

PF 16-JAN-2001; 2001WO-US01574.

XX

PR 14-JAN-2000; 2000US-0483672.

XX

PA (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Panger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MD;
XX WPI; 2001-425873/45.

DR WPI; 2001-425873/45.

XX New polynucleotide encoding a prostate-specific protein, for

PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX Claim 1; Page 359-360; 543pp; English.

XX The present invention describes polynucleotide sequences (I) which encode

CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
CC AAH01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.

XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 33.8%; Score 522; DB 22; Length 537;
Best Local Similarity 100.0%; Pred. No. 8.7e-119;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1025 GGCCGAGTGGGACCAAGGACTTTTCAGGCTTCTTCGATCATCAAGGCTG 1084
DB 1 GGCCGAGTGGGACCAAGGACTTTTCAGGCTTCTTCGATCATCAAGGCTG 60
QY 1085 CCCCTCTGTGTCATCATGATCAGCCTATGAGTGGCAAAAGCTTTCCAGAGGC 1144
DB 61 CCCCTCTGTGTCATCATGATCAGCCTATGAGTGGCAAAAGCTTTCCAGAGGC 120
QY 1145 TGAACACGACCGGCTTCTGGGGCTGAAAGGGGCAAGGACCCCGTCTCTC 1204
DB 121 TGAACACGACCGGCTTCTGGGGCTGAAAGGGGCAAGGACCCCGTCTCTC 180
QY 1205 CCACGATGGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1264
DB 181 CCACGATGGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 1265 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGAGGAGGAGGAGGAGG 1324
DB 241 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGAGGAGGAGGAGGAGG 300
QY 1325 GGCGCCAGCAGCTTCTTCAGACAACTTCTTCCTGCTCCAGTGGGAGATCA 1384
DB 301 GGCGCCAGCAGCTTCTTCAGACAACTTCTTCCTGCTCCAGTGGGAGATCA 360
QY 1385 CTTACACACCCCAAGTCTCAAGCAAAATCTTCAGTCCGCCCTTCGTTTCCCTGT 1444
DB 361 CTTACACACCCCAAGTCTCAAGCAAAATCTTCAGTCCGCCCTTCGTTTCCCTGT 420
QY 1445 GTTCTCTAGCTGGGATGCTCTCAGGAGCCAGAGAGCCCTCAGCCTGTAGTCTCC 1504
DB 421 GTTCTCTAGCTGGGATGCTCTCAGGAGCCAGAGAGCCCTCAGCCTGTAGTCTCC 480
QY 1505 CTGACCTTGTAAATCTTAACTCTTAAAGATGATGAATTC 1546
DB 481 CTGACCTTGTAAATCTTAACTCTTAAAGATGATGAATTC 522

RESULT 8

AAH85036

ID AAH85036 standard; cDNA; 537 BP.

XX AC AAH85036;

XX DT 25-SEP-2001 (first entry)

XX Human prostate-specific cDNA sequence CGI-69.

DE Human; prostate cancer; therapy; diagnosis; cat eye syndrome;

XX Human; prostate cancer; therapy; diagnosis; cat eye syndrome;

XX Chromosome 22q11.2; prostate-specific protein; chromosome 1;

XX Prostate specific antigen; PSA; ss.

XX Homo sapiens.

XX WO200134802-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30904.

XX 12-NOV-1999; 99US-0439313.

XX 18-NOV-1999; 99US-043686.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;

PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;

XX WPI; 2001-308785/32.

XX Isolated polypeptide comprising at least an immunogenic portion of a

PT prostate-specific protein, useful in the diagnosis and therapy of a

PT prostate cancer -

XX Claim 3; Page 259; 325pp; English.

XX The present invention describes an isolated polypeptide (PI) comprising

CC at least an immunogenic portion of a prostate-specific protein, or its

CC variant. Also described are polynucleotides (NI) encoding (PI). (PI) and

CC (NI) have cytostatic activity and can be used in vaccine production.

CC The polypeptides, nucleic acids and antibodies from the present

CC invention are useful in the diagnosis and therapy of prostate cancer.

CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located

CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome

CC region. Prostate specific antigen (PSA) P501S was located on

CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent

CC polynucleotide and polypeptide sequences used in the exemplification

CC of the present invention.

XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 33.8%; Score 522; DB 22; Length 537;

Best Local Similarity 100.0%; Pred. No. 8.7e-119;

Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGCCGAGTGGGACCAAGGACTTTTCAGGCTTCTTCGATCATCAAGGCTG 1084

DB 1 GGCCGAGTGGGACCAAGGACTTTTCAGGCTTCTTCGATCATCAAGGCTG 60

QY 1085 CCCCTCTGTGTCATCATGATCAGCCTATGAGTGGCAAAAGCTTTCCAGAGGC 1144

DB 61 CCCCTCTGTGTCATCATGATCAGCCTATGAGTGGCAAAAGCTTTCCAGAGGC 120

QY 1145 TGAACACGACCGGCTTCTGGGGCTGAAAGGGGCAAGGACCCCGTCTCTC 1204

DB 121 TGAACACGACCGGCTTCTGGGGCTGAAAGGGGCAAGGACCCCGTCTCTC 180

QY 1205 CCACGATGGGAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1264

DB 181 CCACGATGGGAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240

QY 1265 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGAGGAGGAGGAGGAGG 1324

PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX (XUJ)/ XU J.
 PA (DILL)/ DILLON D C.
 PA (MITC)/ MITCHAM J L.
 PA (HARL)/ HARLOCKER S L.
 PA (JIAN)/ JIANG Y.
 PA (KALO)/ KALOS M D.
 PA (FANG)/ FANGER G R.
 PA (RETT)/ RETTER M W.
 PA (STOL)/ STOLK J A.
 PA (DAYC)/ DAY C H.
 PA (VEDV)/ VEDVICK T S.
 PA (CART)/ CARTER D.
 PA (LISX)/ LI S X.
 PA (WANG)/ WANG A.
 PA (SKEL)/ SKEIKY Y A W.
 PA (HEPL)/ HEPLER W T.
 PA (HEND)/ HENDERSON R A.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX WPI; 2002-255649/30.
 DR
 XX New prostate-specific polynucleotides for diagnosing and treating
 PT diseases, in particular prostate cancer, and as markers for the
 PT progression of cancer -
 XX
 PS Claim 1: SEQ ID NO 387; 87pp; English.
 XX
 CC The present invention provides prostate-specific coding sequences and
 CC their encoded proteins. These can be used in the diagnosis and treatment
 CC of cancers, particularly prostate cancer. The present sequence is a cDNA
 CC described in the invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 PS
 PS Query Match 33.8%; Score 522; DB 24; Length 537;
 PS Best Local Similarity 100.0%; Pred. No. 8.7e-119;
 PS Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1025 GGGCCGAGTCGGGACCAAGGACTCTTTGACAGGCTTCCTTCGATCATCAGGCTG 1084
 Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTGACAGGCTTCCTTCGATCATCAGGCTG 60
 QY 1085 CCCCTCTGTCGATCATGATCAGCACTATGATTCGCAAAAGCTTCCAGAGGC 1144
 Db 61 CCCCTCTGTCGATCATGATCAGCACTATGATTCGCAAAAGCTTCCAGAGGC 120
 QY 1145 TGAAACAGGACCGCTTCTGGGCGCTGAAAGGGGCAAGAGGACCCCGTCTCTC 1204
 Db 121 TGAAACAGGACCGCTTCTGGGCGCTGAAAGGGGCAAGAGGACCCCGTCTCTC 180
 QY 1205 CCACGGATGGGAGGAGGAGGAGGACCCAGGACCAAGTCCAGGCGAGGCTGCTCTGG 1264
 Db 181 CCACGGATGGGAGGAGGAGGAGGAGGACCCAGGACCAAGTCCAGGCGAGGCTGCTCTGG 240
 QY 1265 GAGGGGGCTGTTTCCCTTCCTCCCGGCGACAAAGCTCCAGGCGAGGCTGCTCTGG 1324
 Db 241 GAGGGGGCTGTTTCCCTTCCTCCCGGCGACAAAGCTCCAGGCGAGGCTGCTCTGG 300
 QY 1325 GGGGGCCAGCACTTCTCAGACACAACTTCTCTGCTGCTCCAGTCGTGGGATCATCA 1384
 Db 301 GGGGGCCAGCACTTCTCAGACACAACTTCTCTGCTGCTCCAGTCGTGGGATCATCA 360

QY 1385 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTCCGCCCTTCGTGTTCCCTGT 1444
 Db 361 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTCCGCCCTTCGTGTTCCCTGT 420
 QY 1445 GTTGTGCTAGCTGGCAGTGTCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 1504
 Db 421 GTTGTGCTAGCTGGCAGTGTCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 480
 QY 1505 CTGACCCCTGTTAAATTCCTTAAGTCTTAAGATGAATGAACTTC 1546
 Db 481 CTGACCCCTGTTAAATTCCTTAAGTCTTAAGATGAATGAACTTC 522
 RESULT 11
 ABK34735
 ID ABK34735 standard; cDNA; 2170 BP.
 XX
 AC ABK34735;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA for novel secreted protein, SEQ ID 504.
 XX
 KW Human; ss; gene; secreted protein; immune deficiency; viral infection;
 KW bacterial infection; fungal infection; autoimmune disorder; burn;
 KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
 KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
 KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
 KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
 KW lymphoid cell deficiency.
 OS
 OS Homo sapiens.
 XX
 PN WO200177290-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US10295.
 XX
 PR 06-APR-2000; 2000US-194941P.
 XX
 FA (GEMY) GENETICS INST INC.
 XX
 PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;
 XX WPI; 2002-179323/23.
 XX
 PT Six hundred and twenty five polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for
 PT treating immune deficiencies and disorders such as autoimmune disorders
 PT -
 PS Claim 1; Page 269-270; 339pp; English.
 XX
 CC The invention relates to 625 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins, their complements and sequences that hybridise to them.
 CC Also included are a vector comprising the polynucleotide, a host cell
 CC transformed with the vector, the proteins encoded by the
 CC polynucleotides, antibodies that bind to the proteins and identification
 CC of modulators of the proteins or the expression of the polynucleotide.
 CC The polynucleotides can be used as probes for the identification
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 CC and proteins can also be used as nutritional supplements. The protein
 CC is useful in the treatment of various immune deficiencies and disorders
 CC such as viral infections, bacterial infections, fungal infections,
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 CC autoimmune thyroiditis and diabetes) and allergic reactions and
 CC conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's

CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment
 CC of burns, incisions and ulcers. The proteins are also useful for
 CC regulating haematopoiesis, for treating myeloid or lymphoid cell
 CC deficiencies. The present sequence is one of the 625 cDNA sequences
 CC encoding a secreted protein.
 XX

SQ Sequence 2170 BP; 407 A; 663 C; 636 G; 460 T; 4 other;

Query Match 33.5%; Score 517.2; DB 24; Length 2170;
 Best Local Similarity 64.0%; Pred. No. 2e-117;
 Matches 1290; Conservative 0; Mismatches 8; Indels 718; Gaps 5;

QY 100 CAGGACTGAAGCTTCAAGATGGCTGACAGGACCTTGGGGCATCAGCCCTCCAGCAA 159
 DB 148 CAGGACTGAAGCTTCAAGATGGCTGACAGGACCTTGGGGCATCAGCCCTCCAGCAA 207
 QY 160 ATGTGGCTCTCAGGACCGGGCTGTGTACTCTCTCTTCATGACACCCCTGGAGTG 219
 DY 208 ATGTGGCTCTCAGGACCGGGCTGTGTACTCTCTCTTCATGACACCCCTGGAGTG 267
 QY 220 GTGAAGTTTGGCTCTCAGGACCTTGGGGCATCAGCCCTCCAGCAA 279
 DB 268 GTGAAGTTTGGCTCTCAGGACCTTGGGGCATCAGCCCTCCAGCAA 327
 QY 280 AGACTGTGGAGCTCTCTATACCAATGGAAGTGGCTCTCTGTATTCAGTGTGTCTG 339
 DB 328 AGACTGTGGAGCTCTCTATACCAATGGAAGTGGCTCTCTGTATTCAGTGTGTCTG 387
 QY 340 GAGCCTCTGTACTGTGCTCCCAATGTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 399
 DB 388 GAGCCTCTGTACTGTGCTCCCAATGTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 447
 QY 400 CGTTTACTGTGGACATGATGCTCTCTGTAAGATGTAAGTGTGTAAGTGTGTAAGTGTG 459
 DB 448 CGTTTACTGTGGACATGATGCTCTCTGTAAGATGTAAGTGTGTAAGTGTGTAAGTGTG 507
 QY 460 CTCTGGAGCGGCTCTCCCGCACCTCTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCT 519
 DB 508 CTCTGGAGCGGCTCTCCCGCACCTCTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCT 567
 QY 520 ACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
 DB 568 ACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 627
 QY 580 GCACCATGT 610
 D 628 GCACCATGT 687
 QY 611 ----- 610
 DB 688 CTCTGTCTCCCACTCTGGGCCCCCAACCCAGTCAGTCCAGCTCCAGCTCCAGCTCCAGCT 747
 QY 611 ----- 610
 DB 748 CCCTAGGATAGGGGGGTGTTCTAGGTGCCAGACCTGGGACAGCGGTAGGTAGTGA 807
 QY 611 -----TGGGCA 616
 DB 808 TGTAAAGACCTGTGATCTCACTTCTCCCACTGTGTCACCCCAATCCCCCACACAGTGGGCA 867
 QY 617 CCCTGATGTGATAGCCCTCTGAGATGTATGGGACCAAGTGTGAGCTCAGCATGTGT 676
 DB 868 CCCTGATGTGATAGCCCTCTGAGATGTATGGGACCAAGTGTGAGCTCAGCATGTGT 927
 QY 677 CGTACCGGAGCTGGGTGCTGT 736
 DB 928 CGTACCGGAGCTGGGTGCTGT 987
 QY 737 TGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 782
 DB 988 TGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1047

QY 783 ----- 782
 DB 1048 GCTGCAAGGGTGGGTGAGGAAACCTTGGGTGATGACAAAGGGTCTACAGCCNANGTC 1107
 QY 783 ----- 782
 DB 1108 CAANTTTGTGACTAGTGACCATGAAGTTACTTAGCTTTCTTGACCCATGGGTTTCTCAT 1167
 QY 783 ----- 782
 DB 1168 CTGTGATTTGGGCTTAGCAGGTAGAGTATTTGGCCAGNCGGTCAGAGATTAATTCCTAC 1227
 QY 783 ----- 782
 DB 1228 AGAGTGCTGTCTTCTGAGCTGCTATGATAGCAGACAGACTCAGCTGGGAGGAGTGGT 1287
 QY 783 ----- 782
 DB 1288 GGCAGTGAAGGTGGACCGTCTGGGGCCACTTCTTGGCCCTTCCCGTCTCTAATCAGTGC 1347
 QY 783 -----AGCCCTGTACTGTGTTCAACTATGAGCTGGTGAAGAGCTGGTCAATG 829
 DB 1348 ATCCCCCAACCCCAAGCCCTGTACTGTGTTCAACTATGAGCTGGTGAAGAGCTGGTCAATG 1407
 QY 830 GGTTCAGGCTCGAAGAGACAGACTTCTCTGTGGCATGAGCTTTGTGGCTGTGGCATCTCAG 889
 DB 1408 GGTTCAGGCTCGAAGAGACAGACTTCTCTGTGGCATGAGCTTTGTGGCTGTGGCATCTCAG 1467
 QY 890 GGAC ----- 893
 DB 1468 GGCAGCTGAGTGGACTGGGCTGAAGGTGAGGCCAGCGAGGTCTCCCTCGGGGGGGCTG 1527
 QY 894 ----- 893
 DB 1528 AGGATTTTGGAGATGCCAAGAGCAGACCCCATGAGGGCTTTGTCTCCCTCCAGCCGCAA 1587
 QY 894 ----- 893
 DB 1588 GAGCCATCTGGAGGTAGGGGTGGGGTCCAGGAAGTGGAGCTGAGCCCTCTCCCGGCC 1647
 QY 894 -----GGTGGCTGCAAGTCTGACTCTACCTTTCAGCTGGTAAAGACCCAAACGCCA 944
 DB 1648 ACCCCACAGGTGGCTGCAAGTCTGACTCTACCTTTCAGCTGGTAAAGACCCAAACGCCA 1707
 QY 945 GGTGCTCTGGAGCGATGGAGCTGTGAGAGTGAACCCCTGCACTGTGGATCCACCTG 1004
 DB 1708 GGTGCTCTGGAGCGATGGAGCTGTGAGAGTGAACCCCTGCACTGTGGATCCACCTG 1767
 QY 1005 GCTGCTGTGCGAGGATCCGGGCGGAGTCCGGGCAACCAAGGAGCTCTTTGAGGCTTTCCT 1064
 DB 1768 GCTGCTGTGCGAGGATCCGGGCGGAGTCCGGGCAACCAAGGAGCTCTTTGAGGCTTTCCT 1827
 QY 1065 TCCTCGGATCATCAGGCTGCCCTCTCTGTGCACTGATGATGATGATGATGATGATGATGAT 1124
 DB 1828 TCCTCGGATCATCAGGCTGCCCTCTCTGTGCACTGATGATGATGATGATGATGATGATGAT 1887
 QY 1125 CAAAAGCTTTCTTCAGAGCTGAACAGGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGG 1184
 DB 1888 CAAAAGCTTTCTTCAGAGGCTGAACAGGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGG 1947
 QY 1185 AGGCAAGG--ACCCGCTCTCTCCACGATGGGAGA--GGGCAAGGAGAGACCCAGGCCAA 1241
 DB 1948 AGGCAAGGACACCCCTCTCTATCCCAAGATGGGAGATGGGAGAGAGAGACCCAGGCCAA 2007
 QY 1242 GTGCTTTTCTCAGCACTGAGGAGGGGCTTGTGTTTCCCTTCCCTCCCGGCGCAAGCT 1301
 DB 2008 GTGCTTTTCTCAGCACTGAGGAGGGGCTTGTGTTTCCCTTCCCTCCCGGCGCAAGCT 2067
 QY 1302 CCAGGCAAGGCTGTCCCTCTGGGCGGCCAGCACTTCTCTAGACACAACTTCTTCTGTC 1361
 DB 2068 CCAGGCAAGGCTGTCCCTCTGGGCGGCCAGCACTTCTCTAGACACAACTTCTTCTGTC 2127

QY 1362 TGCTCCAGTGGTGGGATCATCACTTACCACCC 1397
 Db 2128 TGCTCCAGTGGTGGGATCATCACTTACCACCC 2163

RESULT 12

ABI99647
 ID ABI99647 standard; cDNA; 704 BP.
 XX AC ABI99647;

XX DT 07-MAR-2002 (first entry)

XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:675.

XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 XX KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

XX OS Mus musculus.

XX PN WO20018188-A2.

XX PD 22-NOV-2001.

XX PF 18-MAY-2001; 2001WO-JP04192.

XX PR 18-MAY-2000; 2000JP-0145977.

XX PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX DR WPI; 2002-034733/04.

XX PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes -

XX PS Claim 2; Page 1679-1680; 2690pp; English.

XX CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding
 CC the protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.

XX SQ Sequence 704 BP; 132 A; 204 C; 209 G; 159 T; 0 other;

Query Match 30.4%; Score 470.4; DB 24; Length 704;
 Best Local Similarity 83.1%; Pred. No. 4.8e-106;
 Matches 560; Conservative 0; Mismatches 111; Indels 3; Gaps 2;

QY 512 TCTACTTCACTGATACCACTGAAGCCCTTCTGTGTGTCGAGCCCTGACCTCTG 571

Db 3 TCGATCCATGGCTTACGACCACTCAAGCCCTTCTGTGTGTCGAGCTGACCTCTG 62

QY 572 ACCTTACGACCCATGTGTGCTGGCGCTGGCCCGCTGGGACCGCTGACTGTGATCA 631

Db 63 ACCTTACGACCCATGTGTGCTGGCGCTGGCCCGCTGGGACCGCTGACTGTGATCA 120

QY 632 GCCCCTGGAGCTTATCGGACAAAGCTGCGAGCTCAGAGCTGCTGACGGAGCTGG 691

Db 121 GCCCCTGGAGCTGCTGGGACCAAGCTGCGAGCTCAGAGCTGCTGATACCGTGAGCTGG 180

QY 692 GTGCTGTGTTGAACTGCACTGCTCAGGGTGGCTGGCGCTCACTGTGGTGGGTGGG 751
 Db 181 CTTCCCTGTTCAGCTCGGTGACTCAGGGTGGCTGGCGCTCTCTGTGGTGGGTGGG 240
 QY 752 GCCCACTGCCCTTCAGAGATGCCCTTCTCAGCCCTCTACTGTTCAATATGAGCTGG 811
 Db 241 GTCCCACTGCTTCAGAGATGCCCTTCTCAGCTCTGACTGTTCAATATGAGCTGG 300
 QY 812 TGAAGAGCTGGCTCAATGGGTTAGGCCGAGAGACCAAGACTTCTGTGGGATGAGCTTTG 871
 Db 301 TGAAGAGCTGGCTGAGCGACTGAGACCAAGAGACCAAGACTTCTGTGGGATGAGCTTTG 871
 QY 872 TGCTGTGTGCTATCTCAGGAGCGGTGCTGCACTGCTGAGCTTACCTTTGAGCTGGTAA 931
 Db 361 TGCGCGGTGG-ATCTCAGGAATGCTGCTGCCACCTTACCTACCTTCCGATGTGGTGA 419
 QY 932 AGACCCAGCGCAGGTGCTCTGGAGCGATGAGGCTGTGAGAGTGAACCCCTTGCATG 991
 Db 420 AGACACAGCGACAGATGTCACTGGAGCGGTGGAGGCTGTGAGAGTGAACCCCTTGCATG 479
 QY 992 TGGACTCCACTTGGCTGCTGCTGGAGGATCCGGGCGAGTGGGACCAAGGACTCT 1051
 Db 480 TGAATCTCAGCTGGCTCTCTGCGAGAAATCCGGGCTGAATCTGGCACCAGGACTCT 539
 QY 1052 TTGCAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCCCTCTCTGTCATCATGATCAGCA 1111
 Db 540 TTGCAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCCCTCTCTGTCATCATGATCAGCA 599
 QY 1112 CTTATGAGTTCGGCAAAAGCTTCTTCCAGAGCTGAACCCAGGACCGGCTTCTGGGCGGCT 1171
 Db 600 CTTATGAGTTCGGCAAAAGCTTCTTCCAGAGCTGAACCCAGGACCGGCTTCTGGGCGGCT 659
 QY 1172 GAAAGCGCAAGCA 1185
 Db 660 TGAAGTCCGGAGCA 673

RESULT 13

ABN94635/c

ID ABN94635 standard; DNA; 452 BP.

AC ABN94635;

XX AC ABN94635;

DT 13-AUG-2002 (first entry)

XX DE Gene #1133 used to diagnose liver cancer.

XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN WO200229103-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US30589.

XX PR 02-OCT-2000; 2000US-237054P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX DR WPI; 2002-426119/45.

XX PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -

Claim 1: SEO ID NO 1133: 298pp; English.

Claim 1; SEQ ID NO 1133; 29pp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytotactic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences.

XX
SQ Sequence 452 BP; 101 A; 117 C; 139 G; 95 T; 0 other;

```

y Match      28.8%; Score 445.4; DB 24; Length 452;
best Local Similarity 99.8%; Pred. No. 5.9e-100;
Matches 446; Conservative 1; Indels 0; Gaps 0;

```

QY	1100	TCATGATCAGCACTATGAGTTGGCAAAAGCTTCTTCAGAGGCTGAACCAAGACCGGC	1159
Db	452	TCATGATCAGCACTATGAGTTGGCAAAAGCTTCTTCAGAGGCTGAACCAAGACCGGC	393
QY	1160	TTCTGGGGCGCTGAAGAGGCGAAGGAGCAGAGACCCCGTCTCTCCACGATATGGGAGA	1219
Db	392	TTCTGGGGCGCTGAAGAGGCGAAGGAGCAGAGACCCCGTCTCTCCACGATATGGGAGA	333
QY	1220	GCGCAGGAGGAGACCCACAGCCAAAGTCCCTTCCTCAGCACTGAGGAGGGGCGTGTTC	1279
Db	332	GCGCAGGAGGAGACCCACAGCCAAAGTCCCTTCCTCAGCACTGAGGAGGGGCGTGTTC	273
QY	1280	CTTTCCTCCGGGGAACAAGCTCCAGGCGAGGGTGTCCCTCTGGGCGGCCAGCACTTC	1339
Db	272	CTTTCCTCCGGGGAACAAGCTCCAGGCGAGGGTGTCCCTCTGGGCGGCCAGCACTTC	213
QY	1340	CTCAGACAACTTCTTCCTGCTGCTCCAGTGTGGGGATCATCATTACCAACCCGCCA	1399
Db	212	CTCAGACAACTTCTTCCTGCTGCTCCAGTGTGGGGATCATCATTACCAACCCGCCA	153
QY	1400	AGTTCAGACCAAAATCTTTCAGCTGCGCCCTTCGTGTTCCCTGTGCTGTAGCTGG	1459
Db	152	AGTTCAGACCAAAATCTTTCAGCTGCGCCCTTCGTGTTCCCTGTGCTGTAGCTGG	93
C	1460	GCATGCTCCAGGACCAAGACGCTCAGCTGGGTAGTCTCCCTGACCTTGTGTAAT	1519
Db	92	GCATGCTCCAGGACCAAGACGCTCAGCTGGGTAGTCTCCCTGACCTTGTGTAAT	33
QY	1520	TCCTTAAGTCTAAAGATGATGAACCTTC	1546
Db	32	TCCTTAAGTCTAAAGATGATGAACCTTC	6

RESULT 14

ABL01419
ID ABL01419 standard: DNA: 758 BP.

XX ABL01419;

XX
DT 15-MAR-2002 (first entry)

XX DE Murine apoptosis related DNA sequence #84.

XX	Apoptosis; mouse; cancer; autoimmune disease; viral infection;
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW	reflexion injury; stroke; liver damage; dilatory cardiomyopathy;
KW	transgenic animal; hepatocarcinoma; antialcoholism; cytotoxic;
KW	immunosuppressive; viricide; neurotropic; neuroprotective; vasotropic;
KW	anti-parkinsonian; cerebrotective; ds.

XX
OS Mus sp.

AA
PN DE10126344-A1.

24-JAN-2002.

XX
PF
30-MAY-2001: 2001DE-1026344.

XX
PP
14-III.-2000. 20000E-1034303.XX
PA (PLAC) MAY BLANK GES FOERDERUNG WISSENSCHAFTEN

DT Crimm S. Schoenfeld N. Brazillis E. Cramer II. Gewies A. Voss E.

PI Mund T, Albayrak T, Gille H, Klein M;
vv

DR WPI; 2002-115563/16.

PT New apoptosis-associated nucleic acid sequences and polypeptides;
PT useful for diagnosis, treatment and prevention of e.g. tumors and
PT neurodegeneration -

PS Claim 1: Page 165: 227pp: German.

The present invention relates to nucleic acids from the mouse, where the nucleic acid is associated with apoptosis. The sequences can be used in the diagnosis, treatment and prevention of diseases associated with excessive or inadequate apoptosis, including tumours, autoimmune diseases, viral infections, degenerative diseases (Alzheimer's disease, Parkinson's and Huntington's diseases), reperfusion injury, stroke and alcohol-induced injury to the liver, for identifying agents for treating these diseases, and to prepare transgenic animals in which expression of an apoptosis related sequence is altered. These are useful for genetic and/or pharmacological investigations of apoptosis and related diseases, including dilatory cardiomyopathy. The present sequence is one of the apoptosis related sequences of the invention.

Sequence 758 BP; 136 A; 244 C; 206 G; 149 T; 23 other; XX SQ

Query Match 28.3%; Score 437.4; DB 24; Length 758;

Query Match 100.0%; Pred. No. 6.5e-98; Indels 27; Gaps 4;
 Best Local Similarity 79.0%; Pred. No. 6.5e-98; Indels 133;
 Matches 565; Conservative 0; Mismatches 123;

34 GGGCGGACGGACGTAATGGGCCCGCCCTGGCGCGCGCCGACGAGCACCAGCC 93

d6 37 GGGCGGAGCCGGTATTGGCCCCGCCAGGCCCTGGCGCGGCACGAACGGCT 96

153

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY Z14 GACGGGAGGGTTCGGCCGACGTCACAGCGGCTCCACATGGCCAGCGCGATGCTT

DB 217 NATGNGGIGAAAGGTCGGCCTTCAGTCTCAGAGAGACCCCTCGGCACACCGAGGATGGACAACT 277

QY 274 TCCTCCAGACTGTGGAGCCCTCTCCCTATACCAAT-----GG 30

Db 277 CCCTCCAGATTCTGGAGTCTCTCCTACACCAATCATCCTCCGCTCTACAGTCCCCAGGG 336

Qy 310 AAGTGCCCTCCTGTATTGCAATGGTGTCTGGAGCCTCTGTACCTGTGCCCAAATGGTGCC 365

Db 337 AAGTGCCCTCTACTGNAATGGAGTCCTGGAGCCCCCTGTACCTGTGCCCAATGGTACC 396

Qy 370 CGCTGTGCCACCTGGTTTCAAGACCCCTACCCGCTTCACTGGCACCATGGATGCCCTTCGTG 429

Ddb 397 CGNTGTGCCACCTGNTTTCAGGACCCACACGGTTCACCTGGCACCTTGATGCCTTNGG 456

100

Dd	457	AAAGTTGNGCGGCGATGAGGCGACTAGGACCCCTGNNANNGCCTTCCCTGAGCCACCCCTGGNG	516
Qy	490	ATGACTGTGCCAGCTACCGCCATCTACTTCACTGCCCTATGACCAACTGAAGGCTTTCCTG	549
Dd	517	ATGACCGNGCAGCTACTGCTATCTACTTCACTGTCTAGCACCAACTCAAGGCTTTCCTG	576
Qy	550	TGTGGTTCGAGCCCTGACCTCTGACCTCTAGCGCCCATGTGGTGGCGGCTGGCCCCG	609
Dd	577	TGTGG-CAGTCTCTTGAAGCTCTGACCTCTAGCCCTTACGCCCATGTGGTGGCTGCCCTGCCCGA	635
Qy	610	CTGGGACCGGAGCTGATGATGAGCCCTGAGCCCTGAGCGTTATCGCGCAAAAGCTGCGAGCTAG	669
Dd	636	ATGGGACCGGTGACGT-TGCAGCCCTTGNAGCTCTGTCGCGCAACAGCTGCGAGCTAG	694
Qy	670	CATGTGTCTGTTACCGGAGCTGGGTGCCCTGTGTTCGAACCTGCAGTGGCTCAGGGTG	724
Dd	695	CATGTGTCTATACCGTGGCTGCTTCTGTGTT-AAGCTGCGNGACTCAGGGTG	748
RESULT	15		
AACT76691/c	ID	AACT76691 standard; cDNA; 6712 BP.	
XX	AC	AACT76691;	
XX	XX		
DT	08-FEB-2001	(first entry)	
XX	XX		
DE	Human ORFX ORF2246	polynucleotide sequence SEQ ID NO:4491.	
XX	XX		
KW	Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;		
KW	vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;		
KW	anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardian;		
KW	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;		
KW	hypotensive; dermatological; immunosuppressive; antiinflammatory;		
KW	antiviral; antibacterial; antifungal; antirheumatic; antithyroid;		
KW	antianemic; gene therapy; cancer; proliferative disorder; hypertension;		
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;		
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;		
KW	cholesterol ester storage; systemic lupus erythematosus; infection;		
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;		
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;		
KW	bone damage; cartilage damage; ant inflammatory disease; coagulation;		
KW	thrombosis; contraceptive; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
PN	WO200058473-A2.		
PD	05-OCT-2000.		
XX	XX		
PF	31-MAR-2000; 2000WO-US08621.		
XX	XX		
PR	31-MAR-1999; 99US-0127607.		
PR	02-APR-1999; 99US-0127636.		
PR	05-APR-1999; 99US-0127728.		
PR	30-MAR-2000; 2000US-0540761.		
XX	XX		
PA	{CURA-} CURAGEN CORP.		
XX	XX		
ZI	Shinkets RA, Leach M;		
XX	XX		
DR	WPI; 2000-602362/57.		
DR	P-PSDB; AAB42482.		
XX	XX		
PT	Novel nucleic acids and peptides derived from open reading frame X,		
PT	useful for treating e.g. cancers, proliferative disorders,		
PT	neurodegenerative disorders and cardiovascular disease -		
XX	XX		
PS	Claim 5; Page 3671-3675; 5507pp; English.		
XX	XX		
CC	AACT74446 to AACT7606 encode the proteins given in AAB40237 to AAB43397,		
CC	which represent the human ORFX open reading frames 1 to 3161. The ORFX		
CC	sequences have activities such as: cytostatic; hepatotropic; vulnery;		

QY 1022 TCCGGGCGAGTCGGGACCAAGGAGCTTTGCGAGGCTTCCTTCGATCATCAAGG 1081
 Db 2851 TTGTTGCTAAAAATGATTTTCGGGATTATTTTCAGGCTAATTCCTCGCTTAATTAAAA 2792
 QY 1082 CTGCCCCCTCTGTGCGATCATGATCAGCACCTATGAGTTGGCAAAAAGCTTCTTCAGA 1141
 Db 2791 TTGCTCCTGCTGTGCGCAATTATGATCAGTACATATGAAATTGGAAAGGCTTTTTCAGA 2732
 QY 1142 GGCTGAA 1148
 Db 2731 AACAAA 2725

Search completed: January 22, 2003, 19:04:05
 Job time : 387.534 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:32:11 ; Search time 63.3511 Seconds
(without alignments)
7484.038 Million cell updates/sec

Title: US-09-888-358-2
Perfect score: 1546
Sequence: 1 ggctagtgctgagcg.....gtctaagatgatgaattc 1546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Tc number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1443.8	93.4	1643	2	US-08-933-750C-68
2	1443.8	93.4	1643	3	US-09-234-613-68
3	522	33.8	537	4	US-09-605-785-387
4	522	33.8	537	4	US-09-439-313-387
5	522	33.8	537	4	US-09-352-616A-387
6	61	3.9	997	3	US-09-188-930-23
7	61	3.9	1816	3	US-09-188-930-262
8	48.8	3.2	1311	4	US-09-068-140A-9
9	48.8	3.2	5150	4	US-09-068-140A-14
10	43.2	2.8	7218	1	US-08-232-463-14
11	42.2	2.7	1594	2	US-08-933-750C-61
12	42.2	2.7	1594	3	US-09-234-613-61
13	40.8	2.6	1259	3	US-08-961-871-11
14	39.8	2.6	1192	4	US-09-142-565-1
15	39.6	2.6	897	4	US-09-434-288-6
16	39.2	2.5	1177	3	US-08-961-871-9
17	38.4	2.5	34303	2	US-08-735-609-4
18	38.4	2.5	34303	2	US-08-735-609-4
19	38.4	2.5	34303	3	US-09-315-372-4
20	38.4	2.5	34303	3	US-09-244-752-4
21	38.4	2.5	34303	3	US-09-245-497-4
22	38.4	2.5	34303	4	US-09-562-919-4
23	38.4	2.5	34382	2	US-08-374-483-6
24	38.4	2.5	35408	4	US-08-973-334-3
25	38.4	2.5	35408	4	US-09-563-869A-3
26	38.4	2.5	35408	4	US-08-549-489-3
27	38.4	2.5	35935	2	US-08-735-609-1

28	38.4	2.5	35935	2	US-08-735-609-1	Sequence 1, Appl
29	38.4	2.5	35935	3	US-08-379-452-43	Sequence 43, Appl
30	38.4	2.5	35935	3	US-09-315-372-1	Sequence 1, Appl
31	38.4	2.5	35935	3	US-09-244-752-1	Sequence 1, Appl
32	38.4	2.5	35935	3	US-09-245-497-1	Sequence 1, Appl
33	38.4	2.5	35935	3	US-09-409-670-43	Sequence 43, Appl
34	38.4	2.5	35935	4	US-09-562-919-1	Sequence 1, Appl
35	37.4	2.4	2634	1	US-08-196-218-31	Sequence 31, Appl
36	37.4	2.4	2634	1	US-08-681-953-31	Sequence 31, Appl
37	36.8	2.4	8460	1	US-08-469-005A-9	Sequence 9, Appl
38	35.8	2.3	793	4	US-08-998-416-471	Sequence 471, App
39	35.6	2.3	1535	4	US-08-910-973-10	Sequence 10, Appl
40	35.6	2.3	1535	4	US-09-499-227-10	Sequence 3, Appl
41	35.6	2.3	1550	3	US-09-234-332-3	Sequence 15, Appl
42	35.6	2.3	1722	4	US-09-385-028-15	Sequence 13, Appl
43	35.6	2.3	11604	4	US-09-385-028-13	Sequence 1, Appl
44	35.6	2.3	15079	4	US-09-385-028-1	Sequence 48, Appl
45	35.4	2.3	606	2	US-08-432-871C-48	

ALIGNMENTS

RESULT 1
US-08-933-750C-68
; Sequence 68, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNCOAT01

CLONE: 724157
US-08-933-750C-68

Query Match 93.4%; Score 1443.8; DB 2; Length 1643;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	100	CAGGACTGAAGCTTCAAGATGCTGACAGGACCCCTCGGGGATCAGCCCTCCAGCAA	159
DB	187	CAGGACTGAAGCTTCAAGATGCTGACAGGACCCCTCGGGGATCAGCCCTCCAGCAA	246
QY	160	ATGFTGSCCTCAGGACCCGSGCTGTGTTACCTCTCTCTTCATGACACCCCTGAGAGT	219
DB	247	ATGFTGSCCTCAGGACCCGSGCTGTGTTACCTCTCTCTTCATGACACCCCTGAGAGT	306
QY	220	GTGAAGGTTTCGCTGACGCTCAGCGGCCCTCCATGSCCAGGAGCTGATGCTTCCTCC	279
DB	307	GTGAAGGTTTCGCTGACGCTCAGCGGCCCTCCATGSCCAGGAGCTGATGCTTCCTCC	366
QY	280	AGACTGTGAGGCTCTCCCTATACCAAAATGGAAGTGCCTCTGTTATGCAATGGTCTTG	339
DB	367	AGACTGTGAGGCTCTCCCTATACCAAAATGGAAGTGCCTCTGTTATGCAATGGTCTTG	426
QY	340	GAGCCTCTGACTGTGCGCCAAATGGTGCCTGCTGCCACCTGTTTCAAGACCCCTACC	399
DB	427	GAGCCTCTGACTGTGCGCCAAATGGTGCCTGCTGCCACCTGTTTCAAGACCCCTACC	486
QY	400	CGCTTCACCTGGCAGCATGATGCTGTAAGATGCTGAGCAGAGGACCCAGGAC	459
DB	487	CGCTTCACCTGGCAGCATGATGCTGTAAGATGCTGAGCAGAGGACCCAGGAC	546
QY	460	CTCTGAGGCGCTCCCGCCACCTGCTGATGATGCTGCCAGCTACGCCATCTACTTC	519
DB	547	CTCTGAGGCGCTCCCGCCACCTGCTGATGATGCTGCCAGCTACGCCATCTACTTC	606
QY	520	ACTGCTATGACAACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTGTGACCTTAC	579
DB	607	ACTGCTATGACAACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTGTGACCTTAC	666
QY	580	GCACCATGTGCTGGCGCTGCGCCCTGGGACCGTGTGATGATGATGATGATGATGATGAT	639
DB	667	GCACCATGTGCTGGCGCTGCGCCCTGGGACCGTGTGATGATGATGATGATGATGATGAT	726
QY	640	GAGCTTATCGGACAAAGCTCAGGCTCAGCATGCTGTGTACCGGAGCTGGGTGCTGT	699
DB	727	GAGCTTATCGGACAAAGCTCAGGCTCAGCATGCTGTGTACCGGAGCTGGGTGCTGT	786
QY	700	GTTGCACTGCACTGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	759
DB	787	GTTGCACTGCACTGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	846
QY	760	GCCTTCAGAGTGCCTTCTCAGCCCTGTACTGTTCACTATGATGCTGTGAAGAC	819
DB	847	GCCTTCAGAGTGCCTTCTCAGCCCTGTACTGTTCACTATGATGCTGTGAAGAC	906
QY	820	TGGCTCAATGGGTTTCAAGCCCAAGGACCACTTCTGTGGGCTGAGCTTGTGGCTGCT	879
DB	907	TGGCTCAATGGGTTTCAAGCCCAAGGACCACTTCTGTGGGCTGAGCTTGTGGCTGCT	966
QY	880	GCGCTCAGGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	939
DB	967	GCGCTCAGGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1026
QY	940	CGCAGGTCGCTTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCTGCTGCTGCTGCT	999
DB	1027	CGCAGGTCGCTTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCTGCTGCTGCTGCT	1086
QY	1000	ACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1059
DB	1087	ACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1146
QY	1060	TTCTCTCTCGATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1119

DB	1147	TTCTTCTCTCGATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1206
QY	1120	TTCTTCTCTCGATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1179
DB	1207	TTCTTCTCTCGATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1266
QY	1180	CAAGGAGGCAAGGACCCCTGCTCTCCACGATGGGAGAGGAGGAGGAGGAGGAGGAGGAGG	1239
DB	1267	CAAGGAGGCAAGGACCCCTGCTCTCCACGATGGGAGAGGAGGAGGAGGAGGAGGAGGAGG	1326
QY	1240	AAGTGCCTTTTCTCAGCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1299
DB	1327	AAGTGCCTTTTCTCAGCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1386
QY	1300	CTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1359
DB	1387	CTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1446
QY	1360	GCTGCTCAGTCTGCTGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCTTCC	1419
DB	1447	GCTGCTCAGTCTGCTGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCTTCC	1506
QY	1420	AGCTGCCCCCTTCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1479
DB	1507	AGCTGCCCCCTTCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1566
QY	1480	AAGCCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1539
DB	1567	AAGCCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1626
QY	1540	GAACTTC 1546	
DB	1627	GAACTTC 1633	

RESULT 2
US-09-234-613-68
; Sequence 68, Application US/09234613
; Patent No. 613 973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guebler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US

```

,
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: 415-855-0555
,
, TELEFAX: 415-845-4166
,
,
,
,
, INFORMATION FOR SEQ ID NO: 68:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 1643 base pairs
,
, TYPE: nucleic acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, IMMEDIATE SOURCE:
,
, LIBRARY: SYNCOAT01
,
, CLONE: 724157
,
,
, US-09-234-613-68

```

Query Match 93.4%; Score 1443.8; DB 3; Length 1643;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 1445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Q	100	CAGGACTGAAGCTTTCAAGATGGCTGACACAGAGACCTCTGGGGCATCAGCCCCCTCCAGCAA	159
D	187	CAGGACTGAAGCTTTCAAGATGGCTGACACAGAGACCTCTGGGGCATCAGCCCCCTCCAGCAA	246
Q	160	ATGGTGGCCTTCAGGCAACCGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGGAGGTG	219
D	247	ATGGTGGCCTTCAGGCAACCGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGGAGGTG	306
Q	220	GTGAAGGTTTGGCTGCAAGTCTCAGCGGCCCTCCNTGGCAGCAGAGCTGATGCTTCTCTCC	279
D	307	GTGAAGGTTTGGCTGCAAGTCTCAGCGGCCCTCCNTGGCAGCAGAGCTGATGCTTCTCTCC	366
Q	280	AGACTGTGAGGCTCTCTCTATACCAAAATGGAAGTGCCTCTGTATTGCAATGGTGTCTCTG	339
D	367	AGACTGTGAGGCTCTCTCTATACCAAAATGGAAGTGCCTCTGTATTGCAATGGTGTCTCTG	426
Q	340	GAGCCTCTGTACTGTGGGCCAAATGGTGGCCGCTGTGCACCTGTTTCAAGACCCCTACC	399
D	427	GAGCCTCTGTACTGTGGGCCAAATGGTGGCCGCTGTGCACCTGTTTCAAGACCCCTACC	486
Q	400	CGCTTCACTGGCACCATGATGCTCTCGTGAAGATCGGTGAGGCACAGGGACACAGAGCC	459
D	487	CGCTTCACTGGCACCATGATGCTCTCGTGAAGATCGGTGAGGCACAGGGACACAGAGCC	546
Q	460	CTCTGGAGGGGCTCCCGGCCACCCCTGGTGATGACTGTGCAGCTACCGGCATCTACTTC	519
D	547	CTCTGGAGGGGCTCCCGGCCACCCCTGGTGATGACTGTGCAGCTACCGGCATCTACTTC	606
C	520	ACTGCCTATGACCAACTGAAGGCCCTTCTGTGTGGTGGAGCCCTGACCTCTGACCTCTAC	579
D	607	ACTGCCTATGACCAACTGAAGGCCCTTCTGTGTGGTGGAGCCCTGACCTCTGACCTCTAC	666
Q	580	GCACCCATGTTGGCTGGGCGCTGGCCGCTGGGCAACCTGACCTGACTGTGATCAGGCCCTGTG	639
D	667	GCACCCATGTTGGCTGGGCGCTGGGCGCTGGGCAACCTGACCTGACTGTGATCAGGCCCTGTG	726
Q	640	GAGCTTATCGGACAAAGCTGCAGGCTCAGCATGTGTGCTGATCGGAGCTCGGTGCTGT	699
D	727	GAGCTTATCGGACAAAGCTGCAGGCTCAGCATGTGTGCTGATCGGAGCTCGGTGCTGT	786
Q	700	GTTTCGAAGCTGCAGTGGCTTCAGGTTGCTGGCGCTCACTGTGGCTGGGTGGGGCCCCACT	759
D	787	GTTTCGAAGCTGCAGTGGCTTCAGGTTGCTGGCGCTCACTGTGGCTGGGTGGGGCCCCACT	846
Q	760	GCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTNTGAGCTGTGGAAGC	819
D	847	GCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTNTGAGCTGTGGAAGC	906
Q	820	TGGCTCAATGGGTTTCAGCGGAAGCCAGACCTTCTGTGGGCAATGAGCTTTGTGGCTGT	879
D	907	TGGCTCAATGGGTTTCAGCGGAAGCCAGACCTTCTGTGGGCAATGAGCTTTGTGGCTGT	966
Q	880	GGCATCTCAGGAGCGGTGGCTGCAGTGTGACTCTACCTTTGAGCTGTGAAGACCCAA	939

; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-605-785-387

Query Match 33.8%; Score 522; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1025 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCCCTCTGTGCCATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGC 1144
Db 61 CCCCCCTCTGTGCCATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGC 120
QY 1145 TGAACGAGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAGGCTCTCTC 1204
Db 121 TGAACGAGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAGGCTCTCTC 180
QY 1205 CCACGATGGGAGAGGAGGAGAGAGCCAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 1264
Db 181 CCACGATGGGAGAGGAGGAGAGAGCCAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 240
QY 1265 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 1324
Db 241 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 300
QY 1325 GCGGCCAGCACTTCTCAGACACACTTCTTCCTGCTGCTCCAGTCTGGGATCATCA 1384
Db 301 GCGGCCAGCACTTCTCAGACACACTTCTTCCTGCTGCTCCAGTCTGGGATCATCA 360
QY 1385 CTTACCCACCCCCCAAGTTCAGACCAAAATCTTCAGCTGCCCTTCGTTTCCCTGT 1444
Db 361 CTTACCCACCCCCCAAGTTCAGACCAAAATCTTCAGCTGCCCTTCGTTTCCCTGT 420
QY 1445 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAACCTC 1546
Db 481 CTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAACCTC 522

RESULT 4
US-09-439-313-387
; Sequence 387, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Soik, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-439-313-387

Query Match 33.8%; Score 522; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1025 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCCCTCTGTGCCATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGC 1144
Db 61 CCCCCCTCTGTGCCATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGC 120
QY 1145 TGAACGAGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAGGCTCTCTC 1204
Db 121 TGAACGAGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAGGCTCTCTC 180
QY 1205 CCACGATGGGAGAGGAGGAGAGCCAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 1264
Db 181 CCACGATGGGAGAGGAGGAGAGCCAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 240
QY 1265 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 1324
Db 241 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 300
QY 1325 GCGGCCAGCACTTCTCAGACACACTTCTTCCTGCTGCTCCAGTCTGGGATCATCA 1384
Db 301 GCGGCCAGCACTTCTCAGACACACTTCTTCCTGCTGCTCCAGTCTGGGATCATCA 360
QY 1385 CTTACCCACCCCCCAAGTTCAGACCAAAATCTTCAGCTGCCCTTCGTTTCCCTGT 1444
Db 361 CTTACCCACCCCCCAAGTTCAGACCAAAATCTTCAGCTGCCCTTCGTTTCCCTGT 420
QY 1445 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAACCTC 1546
Db 481 CTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAACCTC 522

RESULT 5
US-09-352-616A-387
; Sequence 387, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-352-616A-387

Query Match 33.8%; Score 522; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCAGTTCGGGACCAAGGAGCTTTTGCAGGCTTCTTCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCAGTTCGGGACCAAGGAGCTTTTGCAGGCTTCTTCTCGGATCATCAAGGCTG 60
QY 1085 CCCCCCTCTGTCCATCATGATCAGACCTATGATTCGGGAAAAGCTTTTCCAGAGGC 1144
Db 61 CCCCCCTCTGTGCTATCATGATCAGACCTATGATTCGGGAAAAGCTTTTCCAGAGGC 120
QY 1145 TGAACCAAGGACCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGACCCCGCTCTC 1204
Db 121 TGAACCAAGGACCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGACCCCGCTCTC 180
QY 1205 CCAACGATGGGAGAGGCGCAGGAGAGACCCAGCCAAAGTGCCTTTTCTCAGCATGAGG 1264
Db 181 CCAACGATGGGAGAGGCGCAGGAGAGACCCAGCCAAAGTGCCTTTTCTCAGCATGAGG 240
QY 1265 GAGGGGCTTTTCCCTTCCCTCCCGGACAGCTCCAGGAGGAGGCTTCTCCTCTGG 1324
Db 241 GAGGGGCTTTTCCCTTCCCTCCCGGACAGCTCCAGGAGGAGGCTTCTCCTCTGG 300
QY 325 GGGGCCAGCATCTTCTCAGACACAATCTTCTGCTGCTCCAGTCTGTTGGGATCATCA 1384
Db 301 GGGGCCAGCATCTTCTCAGACACAATCTTCTGCTGCTCCAGTCTGTTGGGATCATCA 360
QY 1385 CTTACCCACCCCAAGTTCAGACCAATCTTCCAGTGCCTCCCTTCTGTTTCCCTGT 1444
Db 361 CTTACCCACCCCAAGTTCAGACCAATCTTCCAGTGCCTCCCTTCTGTTTCCCTGT 420
QY 1445 GTTCTGCTGAGTGGGAGTCTCCAGGACCAAGAGCCCTCAGCTGTGTATCTCC 1504
Db 421 GTTCTGCTGAGTGGGAGTCTCCAGGACCAAGAGCCCTCAGCTGTGTATCTCC 480
QY 1505 CTGACCTTGTAACTTCTTAAGTCTTAAAGATGATGAATTC 1546
Db 481 CTGACCTTGTAACTTCTTAAGTCTTAAAGATGATGAATTC 522

RESULT 6
US-09-188-930-23
; Sequence 23, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE OF INVENTION: and Methods For Their Use
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 997
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-23

Query Match 3.9%; Score 61; DB 3; Length 997;
Best Local Similarity 52.1%; Pred. No. 1.1e-06;
Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 393 CCTTACCGCTTCACTGGCACCATGAGTGCCTTCTGTAAGATCGTGAGGACGAGGGCAC 452
Db 718 CCGCAGCAACAACATGTGCATCTAGTGGATTACACAGATGATTCGAGAAGGGGAGC 777
QY 453 CAGGACCTCTGGAGCGGCTCCCGCCACCTCGTGATGACTGTGCCAGCTACCGCCAT 512
Db 778 CAAGTCACTCTGGCGGGCAACGGCATCAATGTCTCAAAATTTGCCCTCAGTCGGCCAT 837

US-09-188-930-23
; Sequence 23, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE OF INVENTION: and Methods For Their Use
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 997
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-23

Query Match 3.9%; Score 61; DB 3; Length 997;
Best Local Similarity 52.1%; Pred. No. 1.1e-06;
Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 393 CCTTACCGCTTCACTGGCACCATGAGTGCCTTCTGTAAGATCGTGAGGACGAGGGCAC 452
Db 718 CCGCAGCAACAACATGTGCATCTAGTGGATTACACAGATGATTCGAGAAGGGGAGC 777
QY 453 CAGGACCTCTGGAGCGGCTCCCGCCACCTCGTGATGACTGTGCCAGCTACCGCCAT 512
Db 778 CAAGTCACTCTGGCGGGCAACGGCATCAATGTCTCAAAATTTGCCCTCAGTCGGCCAT 837

QY 513 CTACTTCACTGCTATGACCACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTCTGA 572
Db 838 CAATTCATGCGCATATGAGCAGATGAAACGGCTTGTGGTAGTATCAGAGACGCTGAG 897
QY 573 CTTCTAGCACCACCATGCTGGCGCTGGCGCCCTGGGGCACCGCTGACTGTGATCAG 632
Db 898 GATCCACGAAAGGCTTGTGGCAGGCTCTTGGCCGAGGCCATTGCCCAGAGTAGCATCTA 957
QY 633 CCCCCTGGAGCTTATGCGGAC 653
Db 958 CCAATGGAGTTCTGAAGAC 978

RESULT 7
US-09-188-930-262
; Sequence 262, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE OF INVENTION: and Methods For Their Use
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 262
; LENGTH: 1816
; TYPE: DNA
; ORGANISM: Mouse
US-09-188-930-262

Query Match 3.9%; Score 61; DB 3; Length 1816;
Best Local Similarity 52.1%; Pred. No. 1.4e-06;
Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 393 CCTTACCGCTTCACTGGCACCATGAGTGCCTTCTGTAAGATCGTGAGGACGAGGGCAC 452
Db 718 CCGCAGCAACAACATGTGCATCTAGTGGATTACACAGATGATTCGAGAAGGGGAGC 777
QY 453 CAGGACCTCTGGAGCGGCTCCCGCCACCTCGTGATGACTGTGCCAGCTACCGCCAT 512
Db 778 CAAGTCACTCTGGCGGGCAACGGCATCAATGTCTCAAAATTTGCCCTCAGTCGGCCAT 837
QY 513 CTACTTCACTGCTATGACCACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTCTGA 572
Db 838 CAATTCATGCGCATATGAGCAGATGAAACGGCTTGTGGTAGTATCAGAGACGCTGAG 897
QY 573 CTTCTAGCACCACCATGCTGGCGCTGGCGCCCTGGGGCACCGCTGACTGTGATCAG 632
Db 898 GATCCACGAAAGGCTTGTGGCAGGCTCTTGGCCGAGGCCATTGCCCAGAGTAGCATCTA 957
QY 633 CCCCCTGGAGCTTATGCGGAC 653
Db 958 CCAATGGAGTTCTGAAGAC 978

RESULT 8
US-09-068-140A-9
; Sequence 9, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road

CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068,140A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04807
FILING DATE: No. 6281409ember 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: C70237
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Ribes nigrum
STRAIN: Ben Alder
US-09-068-140A-9

Query Match 3.2%; Score 48.8; DB 4; Length 1311;
Best Local Similarity 45.7%; Pred. No. 0.0019;
Matches 253; Conservative 0; Mismatches 292; Indels 9; Gaps 2;

QY 398 CCGCTTCACTGGAGCGGCTCCCGCCACCTCGTGTGATGACTGTGCCAGCTACCGCCATCTACT 517
Db 206 CTCAATCCCGCGGTCCGACAGCCCTTGGGTCGATCTGAAAGTTGAAGTTCGCGCG 265
QY 458 CCCTCTGGAGCGGCTCCCGCCACCTCGTGTGATGACTGTGCCAGCTACCGCCATCTACT 517
Db 266 GACTTTACCGTGGCATTGGTGAATGGTCTCGGTGAGGACCTCAGCAGTGTATT 325
QY 518 TCAGTCCCTATGACCACTGAAGCCTTCTGTGTGTGAGGCGCTGACCTCTGAGCTCT 577
Db 326 TCTCCGTTTACGAGATGTGTAAAGAGACTTTTCTCATGGTGTATCCGAGC---AATTCG 382
QY 578 ACCACCCATGTGTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTGATCAGCCCC 637
Db 383 GTGCGGACCGCTTTCGGGGGTGTTCGGACGGTGGCAGCGGTGATTAACCGCA 442
QY 638 TGGAGCTTATGCGGAAGAAGCTGAGGCTCAGCATGTGTGCTGAGGAGCTGGTGCCT 697
Db 443 TGGATGTGTGAACAGAGGTTGCAAGTTCGAGGAGCTCCGTACAGGGTGTGTTGATT 502
QY 698 GTGTGCAACTGAGTGTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 757
Db 503 GCGTGAAGGAGGTTGTGTGAAGAAGAGGATTTGGCGCATTTTACGCATCTTATCGAACAA 562
QY 758 CTGCGCTTCGAGATGTGCTCTCTGAGCCTGTACTGTGTTCAATGAGCTGCTGAAGA 817
Db 563 CTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 622
QY 818 GTTGGCTCAATGGGTTGAGGCGGAA-----GGACGAGACTTCTGTGGGATGAGCTTTG 871
Db 623 AAGGTTGTGTGAGGTTGTGCGCGGAGACTCGGAACGATGAGAAATTTGTTAGTGCATGCTA 682

QY 872 TGGCTGTGTCATCTCAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931
Db 683 CTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
QY 932 AGACCCAAAGCCAG 945
Db 743 AAACCTCAGTTGCAG 756

RESULT 9
US-09-068-140A-14
Sequence 14, Application US/09068140A
Patent No. 6281409
GENERAL INFORMATION:
APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
APPLICANT: and Rex Michael Brennan
TITLE OF INVENTION: Blackcurrant Promoters and Genes
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068,140A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04807
FILING DATE: No. 6281409ember 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: C70237
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 5150 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Ribes nigrum
STRAIN: Ben Alder
US-09-068-140A-14

Query Match 3.2%; Score 48.8; DB 4; Length 5150;
Best Local Similarity 45.7%; Pred. No. 0.0032;
Matches 253; Conservative 0; Mismatches 292; Indels 9; Gaps 2;
QY 398 CCGCTTCACTGGAGCGGCTCCCGCCACCTCGTGTGATGACTGTGCCAGCTACCGCCATCTACT 457
Db 3370 CTCAATCCCGCGGTCCGACAGCCCTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3429
QY 458 CCCTCTGGAGCGGCTCCCGCCACCTCGTGTGATGACTGTGCCAGCTACCGCCATCTACT 517
Db 3430 GACTTTACCGTGGCATTGGTGCATGGTCTCGGTGAGGACCTCAGCAGTGTATT 3489
QY 518 TCAGTCCCTATGACCACTGAAGGCTTCTGTGTGTGTCGAGCCCTGACCTCTGACCTCT 577

Db 3490 TCTCCGTTACGAGATGTGAAGAGACITTTTCTCATGGTATCCGAGC---AATTCCG 3546
Qy 578 ACCACCCATGTGCTGCGCGCTGCGCGCTGGGACGCTGACTGTGATCAGCCCC 637
Db 3547 GTGCGACCCGTTTTCGGGGGTGTTCGCGAGGTGGCAAGCGGTGATTACGCCGA 3606
Qy 638 TGGAGCTTATGGGACAAAGCTCAGAGCTCAGCATGTGTGCTACCGGAGCTGGTGCCT 697
Db 3607 TGGATGTGTGAACACAGAGTTGAGTTGAGAGAGTCCGTACAGGGTGTGTGATT 3666
Qy 698 GTGTTGAACTCAGTGTGCTCAGGTGGCTGGCGCTCACTGCTGGGCTGGGGCCCCA 757
Db 3667 GCCTGAGGAGGTGTGTGTAGAGAGGATTTGGCGCATTTTACGATCTTATCGAACA 3726
Qy 758 CTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 817
Db 3727 CTGCTGTATGATGCCCGCTTACGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3786
Qy 818 GCTGGCTCAATGGTTCAGGCCGAA-----GGACACAGACTTCTGTGGCATGAGCTTGT 871
Db 3787 AAGGGTGTGTGAGGTGTGCGCGGAGACTGCGAAGATGATGAGATTTGTGATGCTGCTA 3846
Qy 872 TGCTGTGTCATCTCAGGACGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTA 931
Db 3847 CTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3906
Qy 932 AGACCCACGCCAG 945
Db 3907 AACTCAGTTGCAG 3920

RESULT 10
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-P1s
US-08-232-463-14
Query Match 2.8%, Score 43.2; DB 1; Length 7218;
Best Local Similarity 3.4%; Pred. No. 0.1;
Matches 9; Conservative 156; Mismatches 99; Indels 0; Gaps 0;
Qy 1261 GAGGAGGGGGCTTGTTCCTCCCTCCCGGCGACAAAGCTCCAGGGCAGGGCTGTCCT 1320
Db 1052 GAGGAGCTTGCATTT 1111
Qy 1321 CTGGCGGGCCAGCAGCTTCTCAGACACAACTTCTTCTGCTGCTGCCAGTCTGGGGATC 1380
Db 1112 YY 1171
Qy 1381 ATCACTACCAACCCCAAGCTTCAAGACCAATCTTCAGCTGCCCTTCGTTTC 1440
Db 1172 YY 1231
Qy 1441 CTGTGTTTCTGTAGCTGGCATGTCTCCAGGACCAAGAGCCCTCAGCTGTTAGT 1500
Db 1232 YY 1291
Qy 1501 CTCCTGACCTTGTATTCCT 1524
Db 1292 YY 1315

RESULT 11
US-08-933-750C-61
; Sequence 61, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Inocyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

```
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNOT02
; CLONE: 207452
US-08-933-750C-61

Query Match      2.7%; Score 42.2; DB 2; Length 1594;
Best Local Similarity 59.7%; Pred. No. 0.11;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1031 AGTCGGGACCAAGGACTCTTTCAGGCTTCTTCGCGATCATCAAGGCTGCCCT 1090
Db 1001 AGAAGGCCCTCGGCTTCTTCAAGGGCTGTCCCGAGCTTCTGAAGGCTGCCCT 1060
QY 1091 CCTGTCCCATCATCATGACACCTATGATCGGCAAAAGCTTCTCCAGAGGCTGAAC 1149
Db 1061 CCACAGGCTCATGTTCTTCTGTATGAATTTCTGTATATGTTCTTCCACTGATGAAC 1119

RESULT 12
US-09-234-613-61
; Sequence 61, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeci
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNOT02
; CLONE: 207452
US-09-234-613-61

Query Match      2.7%; Score 42.2; DB 3; Length 1594;
Best Local Similarity 59.7%; Pred. No. 0.11;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1031 AGTCGGGACCAAGGACTCTTTCAGGCTTCTTCGCGATCATCAAGGCTGCCCT 1090
Db 1001 AGAAGGCCCTCGGCTTCTTCAAGGGCTGTCCCGAGCTTCTGAAGGCTGCCCT 1060
QY 1091 CCTGTCCCATCATCATGACACCTATGATCGGCAAAAGCTTCTCCAGAGGCTGAAC 1149
Db 1061 CCACAGGCTCATGTTCTTCTGTATGAATTTCTGTATATGTTCTTCCACTGATGAAC 1119

RESULT 13
US-08-961-871-11
; Sequence 11, Application US/08961871
; Patent No. 6013858
; GENERAL INFORMATION:
; APPLICANT: Wallace, Douglas C.
; APPLICANT: Graham, Brett H.
; APPLICANT: MacGregor, Grant R.
; TITLE OF INVENTION: Mouse Lacking Heart-Muscle Adenine
; NUCLEOTIDE TRANSLATOR Protein and Methods
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Wimmer and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,871
; FILING DATE: 31-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,017
; FILING DATE: 01-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..1086
; US-08-961-871-11

Query Match      2.6%; Score 40.8; DB 3; Length 1259;
Best Local Similarity 50.0%; Pred. No. 0.22;
Matches 102; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
```

QY	403	TTCTACTGGCACCATGATGGCTCTCGTGAAGATCGTAGGCACGAGGACGACGAGCACCTC	462
Db	340	TACAAAGGCATCATTTGATTGTCTCGTAGAATCCCAAGGAGCAGGGCTTCTCTCTCTTC	399
QY	463	TGAGAGCGGCTCCCGCCACCTGGTGATGACTGTGCCAGTACCGCCATCTATTCTACT	522
Db	400	TGGAGGGGTAACTGTGCCCAACGTGATCCGGTACTTCCCACTCAAGCCCTGAATTCGCC	459
QY	523	GCCTATGACCAACTGAAGCGCTTCTGTGTGTGTGGAGCCCTGACCTCTGACCTCTACGCA	582
Db	460	TTCAAGACAAAGTACAAACGAGATCTCTGGGAGCGTGGATCGCCATGACGATTTCTGG	519
QY	583	CCCATGTGTGGCTGGCGCGCTGGCC	606
Db	520	CGCTACTTTTCTGTTGAATCACTGGCC	543

RESULT 14

```

US-09-142-565-1
; Hence 1, Application US/09142565A
; Patent No. 6187560
; SERIAL INFORMATION:
; APPLICANT: Lee James Beeley
; APPLICANT: Kelly Paine
; APPLICANT: Robert James
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30002
; CURRENT APPLICATION NUMBER: US/09/142,565A
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 9704551.2
; EARLIER FILING DATE: 1997-03-05
; EARLIER APPLICATION NUMBER: 9705614.7
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: 97305305.1
; EARLIER FILING DATE: 1997-07-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: fast-SEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1192
; TYPE: DNA
; ORGANISM: HOMO SAPIEN
US-09-142-565-1

```

Query Match 2.6%; Score 39.8; DB 4; Length 1192;
Best Local Similarity 46.8%; Pred. No. 0.4;
Matches 196; Conservative 0; Mismatches 217; Indels 6; Gaps 2.

Qy	403	TTCTCTGGCCACCATGGATGCTTCGTGAAGATCGTGAAGCAGAGGGCACACAGACCTTC	462
D1	673	TACAGGGGACTATTGAGCGCTACAGAACCATCCAGGGGAGAGGATCAGGGCCGTG	732
Qy	463	TGAGAGGGCTCCCCGGCACCTTGTTGATGACTGTGCAGCTACCGCACTACTTCTACT	522
Db	733	TGGAAAGGAACATTGCCCAACATCATGAGGAATCTATCGTCAACTGTGTGAGGTGGTG	792
Qy	523	GCCTATGACCAACTGAGGGCTTCCTGTGTGGTGGAGCCCTGACCTCTGACCTCTACGGA	582
Db	793	ACCTACGACATCCTCAGGAGAAAGCTGCTGGACTACACCTCTCTACTGCAAACTTCCCC	852
Qy	583	---CCCAATGTTGGCTGGCGCGCTGGCGCCCTGGGCACAGCTGACTGTGATCAGCCCCCTG	639
Db	853	TGCCACTTGTGCTCTGCTTTGGAGCCGCGCTTCTGTGCCACAGTGGTGGCGCTCCCGCGTG	912
Qy	640	GAGCTTATCGGAGACAAGC---TGCAGGCTCAGCATGTGTCTGTATCCGGAGCTGGGTGCC	696
Db	913	GAGCTGTTGAAGACCCGGTATATGAATCACTCTCAGGCCAGTACTTCACGCCCTCTCAAC	972
Qy	697	TGTGTTTCGAAGCTCAGATGGCTCAGGGTGGCTGGCGTCTACTGTGGCTGGGCTGGGGCCCC	756
Db	973	TGTATGATAAAGATGGTGGCGCCAGGAGGGCCCCACAGCCTTCTTCAAGAGGATTTACACC	1032
Qy	757	ACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGCTGGTGAA	815

D**b** 1033 TCCTTTTTCGTTTGGGATCCTGGAACGTGGTGATGTTCTGTAACCTATGAGCAGCTGAA 1091

RESULT 15

```

US-09-434-288-6
; Sequence 5, Application US/09434288
; Patent No. 6303767
; GENERAL INFORMATION:
; APPLICANT: Betlach C., Melanie
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
; TITLE OF INVENTION: CONSTRUCTS THEREFOR
; FILE REFERENCE: 30062-20030.00
; CURRENT APPLICATION NUMBER: US/09/434.288
; CURRENT FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: 60/107,093
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Streptomyces narbonensis
US-09-434-288-6

```

Query Match	2.6%;	Score 39.6;	DB 4;	Length 897;
Best Local Similarity	47.2%;	Pred. No. 0.4;		
Matches 120;	Conservative 0;	Mismatches 134;	Indels 0;	Gaps 0;

QY	391	GACCTTACCCGGTTCACTGGCACCATGGATGCTTCGTGAAGATCGTGAGGACGAGGGC	450
Db	295	GACCCACACGACCCACCGGGCTCACGGGCTCGTGC CGAGCTCGGACCTGGCGGG	354
QY	451	ACGAGGACCTCTGGAGGGGCTCCCGCCACCCCTGGTGATGACTGTGGCAGCTACCGGC	510
Db	355	ACGGCACCGTGTGTCTGGAGCTCACGACGGGAGGGCGGCGCCGGCTGCTCGCC	414
QY	511	ATCTACTTCACTGCCCTATGACCAACTGAAGCGCTTCTGTGTGGTGAGCGCTTGACTCT	570
Db	415	GGCGTCTCCGAGGACACCGCTCAGCGCGCTCTCCACCTGCCGCCCCCGTCGACTCC	474
QY	571	GACCTCTACGACCCATGGTGGCTGGCGGCTGCCCGCTGGGACCGGTGATGTGATC	630
Db	475	GAGCGGTCTGCCGCCACCGACCCGGACCTGGCCCGCTGTAAACCGGAGGGCCACC	534
QY	631	AGCCCCCTGGAGCT	644
Db	535	GC CGCGTGCACCT	548

Search completed: January 22, 2003, 19:06:08
Job time : 90.3511 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 17:33:27 ; Search time 77.3 Seconds

(without alignments)
8985.296 Million cell updates/sec

Title: US-09-888-358-2

Perfect score: 1546

Sequence: 1 ggctagggtcgctgcgagcg.....gtctaaagatgatgaattc 1546

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Tc number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1546	100.0	1546	10	US-09-888-358-2
2	1443.8	93.4	1643	10	US-09-840-787-68
3	1054.4	68.2	1114	10	US-09-888-358-1
4	522	33.8	537	9	US-10-012-896-387
5	522	33.8	537	9	US-09-895-793-387
6	522	33.8	537	9	US-09-895-814-387
7	522	33.8	537	10	US-09-759-143-387
8	522	33.8	537	10	US-09-780-669-387
9	522	33.8	537	10	US-09-822-827-387
C 10	445.4	28.8	452	10	US-09-880-107-1133
C 11	218.4	14.1	240	10	US-09-880-107-1179
12	185.8	12.0	452	10	US-09-867-701-4486
13	172	11.1	275	10	US-09-908-711-46
C 14	170.2	11.0	512	10	US-09-917-800A-329
15	161	10.4	246	10	US-09-960-352-2543
C 16	74.8	4.8	452	10	US-09-864-761-11674
17	74.8	4.8	496	10	US-09-864-761-11708
18	60	3.9	3334	9	US-09-989-298-288
19	60	3.9	3334	9	US-09-989-293A-288

20	60	3.9	3334	9	US-10-063-547-57	Sequence 57, Appl
21	60	3.9	3334	9	US-09-989-735-288	Sequence 288, App
22	60	3.9	3334	9	US-09-990-444-288	Sequence 288, App
23	60	3.9	3334	9	US-09-989-730-288	Sequence 288, App
24	60	3.9	3334	9	US-09-990-436-288	Sequence 288, App
25	60	3.9	3334	9	US-09-991-181-288	Sequence 288, App
26	60	3.9	3334	9	US-09-993-687-288	Sequence 288, App
27	60	3.9	3334	9	US-09-989-734-288	Sequence 288, App
28	60	3.9	3334	9	US-09-997-653-288	Sequence 288, App
29	60	3.9	3334	9	US-10-063-616-57	Sequence 57, Appl
30	60	3.9	3334	10	US-09-989-722-288	Sequence 288, App
31	60	3.9	3334	10	US-09-989-723-288	Sequence 288, App
32	60	3.9	3334	10	US-09-989-279-288	Sequence 288, App
33	60	3.9	3334	10	US-09-989-727-288	Sequence 288, App
34	60	3.9	3334	10	US-09-989-731-288	Sequence 288, App
35	60	3.9	3334	10	US-09-989-732-288	Sequence 288, App
36	60	3.9	3334	10	US-09-991-073-288	Sequence 288, App
37	60	3.9	3334	10	US-09-980-442-288	Sequence 288, App
38	60	3.9	3334	10	US-09-991-163-288	Sequence 288, App
39	60	3.9	3334	10	US-09-993-604-288	Sequence 288, App
40	60	3.9	3334	10	US-09-990-456-288	Sequence 288, App
41	60	3.9	3334	10	US-09-989-721-288	Sequence 288, App
42	60	3.9	3334	12	US-10-006-867-57	Sequence 57, Appl
C 43	57.6	3.7	144	10	US-09-864-761-28253	Sequence 28253, A
44	55.4	3.6	129	10	US-09-864-761-28284	Sequence 28284, A
45	50.6	3.3	3555	12	US-10-044-090-527	Sequence 527, App

ALIGNMENTS

RESULT 1
US-09-888-358-2
; Sequence 2, Application US/09888358
; Patent No. US20020119137A1
; GENERAL INFORMATION:
; APPLICANT: JEWIN, DAVID
; APPLICANT: ADAMS, SEAN H.
; APPLICANT: YU, XING XIAN
; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 10716/66
; CURRENT APPLICATION NUMBER: US/09/888,358
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,307
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-358-2

Query Match	100.0%	Score 1546;	DB 10;	Length 1546;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1546;	Conservative	0;	Mismatches	0;
			Indels	0;
Gaps				0;
QY	1	GGCTAGGTGCGCTCGGAGCGCGGAGCCACGAGGCGCGACGTAATGGCCCGCCT	60	
Db	1	GGCTAGGTGCGCTCGGAGCGCGGAGCCACGAGGCGCGACGTAATGGCCCGCCT	60	
QY	61	GGCCCTGGGCGCGCGCGGCGGAGCGCGGAGCCACGAGGCGCGACGTAATGGCCCGCCT	120	
Db	61	GGCCCTGGGCGCGCGCGGCGGAGCGCGGAGCCACGAGGCGCGACGTAATGGCCCGCCT	120	
QY	121	GCTGACGAGGACCTTCGCGGCGGAGCGCGGAGCCACGAGGCGCGACGTAATGGCCCGCCT	180	
Db	121	GCTGACGAGGACCTTCGCGGCGGAGCGCGGAGCCACGAGGCGCGACGTAATGGCCCGCCT	180	
QY	181	GCTGTGGTTACTCTCTCTTCATGACACCCCTGAGAGGTTCCGCTGCAGTCT	240	
Db	181	GCTGTGGTTACTCTCTCTTCATGACACCCCTGAGAGGTTCCGCTGCAGTCT	240	

QY 241 CAGCGCCCTCCATGCGCAGCGAGCTGATGCTTCTCCAGAGCTGTGGAGCTCTCCTAT 300
 Db 241 CAGCGCCCTCCATGCGCAGCGAGCTGATGCTTCTCCAGAGCTGTGGAGCTCTCCTAT 300
 QY 301 ACCAAATGGAAGTGCCTCTCTGATTAATGGAATGCTTCTCCAGAGCTGTGGAGCTCTCCTAT 360
 Db 301 ACCAAATGGAAGTGCCTCTCTGATTAATGGAATGCTTCTCCAGAGCTGTGGAGCTCTCCTAT 360
 QY 361 AATGTCGCCGCTGTGGCAGCTGTGTTCAAGACCTTACCGCTTCACTGGCACCATTGAT 420
 Db 361 AATGTCGCCGCTGTGGCAGCTGTGTTCAAGACCTTACCGCTTCACTGGCACCATTGAT 420
 QY 421 GCCTTCGTGAAGATCGTGAAGCAGAGGACACAGAGCCCTCTGAGGCGGCTCCCGCC 480
 Db 421 GCCTTCGTGAAGATCGTGAAGCAGAGGACACAGAGCCCTCTGAGGCGGCTCCCGCC 480
 QY 481 ACCCTGTGATGATGTGCGCAGCTACCGCCATCTACTTCACTGCCCTATGACCAACTGAAG 540
 Db 481 ACCCTGTGATGATGTGCGCAGCTACCGCCATCTACTTCACTGCCCTATGACCAACTGAAG 540
 QY 541 GCCTTCGTGTGTCGAGCCCTGACCTCTGACCTCTAGCGACCCATGCTGTGGCGCG 600
 Db 541 GCCTTCGTGTGTCGAGCCCTGACCTCTGACCTCTAGCGACCCATGCTGTGGCGCG 600
 QY 601 CTGGCCCGCTGGGACCGCTGACTGTGATCAGCCCTCTGAGCTTATGCGGACAAAGCTG 660
 Db 601 CTGGCCCGCTGGGACCGCTGACTGTGATCAGCCCTCTGAGCTTATGCGGACAAAGCTG 660
 QY 661 CAGGCTCAGCATGTGTGCTACCGGAGCTGGTGCCTGTGTTTGAAGCTGAGTGGCTCAG 720
 Db 661 CAGGCTCAGCATGTGTGCTACCGGAGCTGGTGCCTGTGTTTGAAGCTGAGTGGCTCAG 720
 QY 721 GGTGGCTGGCGCTCACTGTGCTGGCTGGGCGCCGCTGCTTCCAGATGTGCCCTTC 780
 Db 721 GGTGGCTGGCGCTCACTGTGCTGGCTGGGCGCCGCTGCTTCCAGATGTGCCCTTC 780
 QY 781 TCAGCCCTGTACTGTGTTCAATATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTCAAGCGG 840
 Db 781 TCAGCCCTGTACTGTGTTCAATATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTCAAGCGG 840
 QY 841 AAGGACAGACTTGTGTGGCAGTGTGTTGGCTGGTGGGCTCTCAGGACGGTGGCT 900
 Db 841 AAGGACAGACTTGTGTGGCAGTGTGTTGGCTGGTGGGCTCTCAGGACGGTGGCT 900
 QY 901 GCAGTGTGCTGACTTACCTTTGAGCTGGTGAAGACCCAGCCAGCTGCTCTGGAGCG 960
 Db 901 GCAGTGTGCTGACTTACCTTTGAGCTGGTGAAGACCCAGCCAGCTGCTCTGGAGCG 960
 QY 961 ATGGAGCTGTGAGAGTGAACCCCTGTCATGTGACTCCACTGGCTGCTGCTGGCGAGG 1020
 Db 961 ATGGAGCTGTGAGAGTGAACCCCTGTCATGTGACTCCACTGGCTGCTGCTGGCGAGG 1020
 QY 1021 ATCCGGCGGAGTGGGACCAAGGAGCTTTTTCAGGCTTCTTCCTCGGATCATCAAG 1080
 Db 1021 ATCCGGCGGAGTGGGACCAAGGAGCTTTTTCAGGCTTCTTCCTCGGATCATCAAG 1080
 QY 1081 GCTGCCCCCTCTGTCATCATGATCAGCACTATGATTTGGCAAAAGCTTCTTCCAG 1140
 Db 1081 GCTGCCCCCTCTGTCATCATGATCAGCACTATGATTTGGCAAAAGCTTCTTCCAG 1140
 QY 1141 AGGCTGAACAGGACCGGCTTCTGGCGGCTGAAGGGGCAAGGAGCAAGGACCCCGTC 1200
 Db 1141 AGGCTGAACAGGACCGGCTTCTGGCGGCTGAAGGGGCAAGGAGCAAGGACCCCGTC 1200
 QY 1201 TCTCCACGATGGGAGAGGAGGAGAGACCCAGGCAAGTGCCTTTTCTTCAGCACT 1260
 Db 1201 TCTCCACGATGGGAGAGGAGGAGAGACCCAGGCAAGTGCCTTTTCTTCAGCACT 1260
 QY 1261 GAGGAGGGGGCTGTTTCTCCCTTCCCTGGGAGCAAGCTTCCAGGCGAGGCTGTCCCT 1320
 Db 1261 GAGGAGGGGGCTGTTTCTCCCTTCCCTGGGAGCAAGCTTCCAGGCGAGGCTGTCCCT 1320
 QY 1321 CTGGCGGCGCCAGCAGCTTCTCTCAGACACAATTTCTTCTGCTGCTCCAGTGTGGGGATC 1380

Db 1321 CTGGCGGCGCCAGCAGCTTCTCTCAGACACAATTTCTTCTGCTGCTCCAGTGTGGGATC 1380
 QY 1381 ATCACTTACCCACCCCAAGTTCAAGACCAATATCTTCCAGTGGCCCTTCTGTTTCC 1440
 Db 1381 ATCACTTACCCACCCCAAGTTCAAGACCAATATCTTCCAGTGGCCCTTCTGTTTCC 1440
 QY 1441 CTGCTTTGCTGTAGTGGGATGTCTCCAGAACCAAGAGCCCTCAGCCCTGTTGTTAGT 1500
 Db 1441 CTGCTTTGCTGTAGTGGGATGTCTCCAGAACCAAGAGCCCTCAGCCCTGTTGTTAGT 1500
 QY 1501 CTCCTGACCCCTTGTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTT 1546
 Db 1501 CTCCTGACCCCTTGTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTT 1546

RESULT 2
 US-09-840-787-68
 ; Sequence 68, Application US/09840787
 ; Patent No. US20020058264A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti
 ; Hillman, Jennifer L.
 ; Bandman, Olga
 ; Shah, Purvi
 ; Au-Young, Janice
 ; Yue, Henry
 ; Guegler, Karl J.
 ; Corley, Neil C.
 ; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
 ; NUMBER OF SEQUENCES: 98
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/840,787
 ; FILING DATE: 23-Apr-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/518,865
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0356 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX: <Unknown>
 ; INFORMATION FOR SEQ ID NO: 68:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1643 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: SYN00AT01
 ; CLONE: 724157
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 68 :
 US-09-840-787-68

Query Match 93.4%; Score 1443.8; DB 10; Length 1643;
 Best Local Similarity 99.9%; Pred No. 0;
 Matches 1445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Q:	100	CAGGACTGAAGCTTCAAGATGGCTGACCAGGACCCCTGGGGCATCAGCCGCTCCAGCAA	159
Db	187	CAGGACTGAAGCTTCAAGATGGCTGACCAGGACCCCTGGGGCATCAGCCGCTCCAGCAA	246
Qy	160	ATGGTGGCTCAGGCACCGGGGTGGTGTACCTCTCTCTCATGACCCGCTGACGCTG	219
Db	247	ATGGTGGCTCAGGCACCGGGGTGGTGTACCTCTCTCTCATGACCCGCTGACGCTG	306
Qy	220	GTGAAGGTTGCGCTGCACTCAGCGGCCCTCCATGCGCAGCAGCTGATGCTTCCTCC	279
Db	307	GTGAAGGTTGCGCTGCACTCAGCGGCCCTCCATGCGCAGCAGCTGATGCTTCCTCC	366
Qy	280	AGACTGTGAGGCTCTCCTATATCAAAATGGAAGTCCCTCTGTTATGCAATGGTGTCTG	339
Db	367	AGACTGTGAGGCTCTCCTATATCAAAATGGAAGTCCCTCTGTTATGCAATGGTGTCTG	426
Qy	340	GAGCCTCTGACTCTGTGCCAAATGGTGGCCGTGTGCCACTCGTGTTCAGAGCCCTACC	399
Dp	427	GAGCCTCTGACTCTGTGCCAAATGGTGGCCGTGTGCCACTCGTGTTCAGAGCCCTACC	486
Qy	400	CGCTTCACTGGCAACATGGATGCTTCTGTGAAGATCGTGAGGCAACGAGGCAACAGGACC	459
Db	487	CGCTTCACTGGCAACATGGATGCTTCTGTGAAGATCGTGAGGCAACGAGGCAACAGGACC	546
Qy	460	CTCTGAGCGGCTCCCGGCCAACCTGGTGATGACTGTGCAGCTACCGCCATCTACTTC	519
Db	547	CTCTGAGCGGCTCCCGGCCAACCTGGTGATGACTGTGCAGCTACCGCCATCTACTTC	606
Qy	520	ACTGCCATATGACCAACTGAAGGCTCTTCTGTGTGTGAGCGCTGACCTGTGACCTCTAC	579
Db	607	ACTGCCATATGACCAACTGAAGGCTCTTCTGTGTGTGAGCGCTGACCTGTGACCTCTAC	666
Qy	580	GCACCATATGTTGGCTGGCGGCTGGGCCGCTGGGCAACGTTGATGATGACGCCCCCTG	639
Db	667	GCACCATATGTTGGCTGGCGGCTGGGCCGCTGGGCAACGTTGATGATGACGCCCCCTG	726
Qy	640	GAGCTTATGCGGACAAGCTGCAGGCTCAGAGCTCAGCATGTCTGTACCGGAGCTGGTCTGT	699
Db	727	GAGCTTATGCGGACAAGCTGCAGGCTCAGAGCTCAGCATGTCTGTACCGGAGCTGGTCTGT	786
Qy	700	GTTTCGAACATGCAGTGGCTCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGGCCCCACT	759
nb	787	GTTTCGAACATGCAGTGGCTCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGGCCCCACT	846
Qy	760	GCCCTTCAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGCTGTGTGAAGAC	819
Dp	847	GCCCTTCAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGCTGTGTGAAGAC	906
Q:	820	TGGCTCAATGGGTTCAGGCCGAAAGGACAGACTCTGTGGSCATGAGCTTTGTGGCTGGT	879
Db	907	TGGCTCAATGGGTTCAGGCCGAAAGGACAGACTCTGTGGSCATGAGCTTTGTGGCTGGT	966
Qy	880	GGCATCTCAGGACCGTGGCTGCAGTGTGACTCTACCTTTTGACGTGGTAAGACCCAA	939
Db	967	GGCATCTCAGGACCGTGGCTGCAGTGTGACTCTACCTTTTGACGTGGTAAGACCCAA	1026
Qy	940	CGCCAGGTGCTCTTGGAGCGATGGAAGCTGTGAGAGTGAACCCCTGCAATGTGACTCC	999
Db	1027	CGCCAGGTGCTCTTGGAGCGATGGAAGCTGTGAGAGTGAACCCCTGCAATGTGACTCC	1086
Qy	1000	ACCTGGCTGCTGCTGCGGAGGATCCGGGCGGAGTGGGCAACAGGAGACTCTTTTCAGGC	1059
Db	1087	ACCTGGCTGCTGCTGCGGAGGATCCGGGCGGAGTGGGCAACAGGAGACTCTTTTCAGGC	1146
Qy	1060	TTCTTCTCTCGAATCATCAAGGCTGCCCTCTGTGCCATCATGATCAGCACTTATGAG	1119
Db	1147	TTCTTCTCTCGAATCATCAAGGCTGCCCTCTGTGCCATCATGATCAGCACTTATGAG	1206
Qy	1120	TTCCGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTCTGGCGGCTGAAAGGGG	1179
Lb	1207	TTCCGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTCTGGCGGCTGAAAGGGG	1266

```

RESULT 3
US-09-888-358-1
; Sequence 1, Application US/09888358
; Patent No. US20020119137A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID
; APPLICANT: ADAMS, SEAN H.
; APPLICANT: YU, XING XIAN
; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 10716766
; CURRENT APPLICATION NUMBER: US/09/888,358
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,307
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-358-1

```

Query Match 68.2%; Score 1054.4; DB 10; Length 1114;
Best Local Similarity 97.8%; Pred. No. 3.9e-274;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

Qy	105	CTGAAGCTTCAAGATGGTGTACAGAGACCTCGGGCATCAGCCGCCCTCCAGCAAAATGGT	164
Db	1	CTGAAGCTTCAAGATGGTGTACAGAGACCTCGGGCATCAGCCGCCCTCCAGCAAAATGGT	60
Qy	165	GGCCTCAGGCACCGGGGCTGTGGTTACCTCTCTCTTCATCACACCCCTGGACGTGGTGAA	224
Db	61	GGCCTCAGGCACCGGGGCTGTGGTTACCTCTCTCTTCATCACACCCCTGGACGTGGTGAA	120
Qy	225	GATTGGCTTCGAGTCTCAGCGGCCCTCCATGACGACGAGCTGATGCTTCTCCACGACT	284
Db	121	GATTGGCTCGAGTCTCAGCGGCCCTCCATGACGACGAGCTGATGCTTCTCCACGACT	180
Qy	285	GTGAGAGCTCTCTTATACCAAT-----GGAAAGTGCCTCTCT	320
Db	181	GTGAGAGCTCTCTATACCAATGTGGCCCTCTCTCCAAATCCAGAGGAAGTGCCTCTCT	240
Qy	321	GTATTGCAATGTGTCTGTGGAGCTGTACTGTGTGCCAAATGTGTGCCCGCTGTGCCAC	380

Db 241 GTATTGCAATGGTGTCTCTGAGGCTCTGTACTGTGCGCAATGGTGCCGGCTGTGCCAC 300
QY 381 CTGGTTTCAAGACCTTACCCGCTTCTACCTGGCAACCATGATGCTCTGTGAAGATCGTGAG 440
Db 301 CTGGTTTCAAGACCTTACCCGCTTCTACCTGGCAACCATGATGCTCTGTGAAGATCGTGAG 360
QY 441 GCACAGGGACACAGGACCTTCTGGAGCGGCTCCCGCCACACCTGGTATGACTGTGCC 500
Db 361 GCACAGGGACACAGGACCTTCTGGAGCGGCTCCCGCCACACCTGGTATGACTGTGCC 420
QY 501 AGCTACCGCATCTACTTCACTGCCTATGACCACTGAAGSCCTTCTGTGTGTGCTGAGC 560
Db 421 AGCTACCGCATCTACTTCACTGCCTATGACCACTGAAGSCCTTCTGTGTGTGCTGAGC 480
QY 561 CCTGACCTCTGACCTCTACGACCCCATGTTGGCTGGCGCTGCGCCGCTGGGACCGT 620
Db 481 CCTGACCTCTGACCTCTACGACCCCATGTTGGCTGGCGCTGCGCCGCTGGGACCGT 540
QY 621 GACTGTGATCAGCCCTGAGCTTATCGGACAAAGCTGAGGCTCAGGCTCAGCATGTGCTA 680
Db 541 GACTGTGATCAGCCCTGAGCTTATCGGACAAAGCTGAGGCTCAGGCTCAGCATGTGCTA 600
QY 681 CCGGAGCTGGTGGCTGTCTCGAACTCCAGTGGCTCAGGCTGGCTGGCTCACTGTG 740
Db 601 CCGGAGCTGGTGGCTGTCTCGAACTCCAGTGGCTCAGGCTGGCTGGCTCACTGTG 660
QY 741 GCTGGCTGGGCGCCCACTGCCCTTCGAGATGTCCTTCTCAGCCCTGTACTGTTCAA 800
Db 661 GCTGGCTGGGCGCCCACTGCCCTTCGAGATGTCCTTCTCAGCCCTGTACTGTTCAA 720
QY 801 CTATCAGCTGTGAGAGCTGGCTCAATGGGTTTCAAGGCTGAGGACGAGCTCTGTGG 860
Db 721 CTATCAGCTGTGAGAGCTGGCTCAATGGGTTTCAAGGCTGAGGACGAGCTCTGTGG 780
QY 861 CATGAGCTTTGTGGCTGGTCACTCAGGACGGTGGCTGAGTGTGACTTCACTCCTT 920
Db 781 CATGAGCTTTGTGGCTGGTCACTCAGGACGGTGGCTGAGTGTGACTTCACTCCTT 840
QY 921 TGAGTGTGTAAGACCAACGACGAGTGGCTCTGGAGCGATGAGGCTGTGAGAGTGA 980
Db 841 TGAGTGTGTAAGACCAACGACGAGTGGCTCTGGAGCGATGAGGCTGTGAGAGTGA 900
QY 981 CCCCCTGATGTGAGCTTCACTGGCTGCTGCTGCGAGGATCGGCGGAGTTCGGGAC 1040
Db 901 CCCCCTGATGTGAGCTTCACTGGCTGCTGCTGCGAGGATCGGCGGAGTTCGGGAC 960
QY 1041 CAAGGAGCTTTTTCAGGCTTCTTCTCGGATCATCAAGGCTGCCCTTCTGTGGCAT 1100
Db 961 CAAGGAGCTTTTTCAGGCTTCTTCTCGGATCATCAAGGCTGCCCTTCTGTGGCAT 1020
QY 1101 CATGATGAGCACTTATGAGTTTGGGAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCT 1160
Db 1021 CATGATGAGCACTTATGAGTTTGGGAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCT 1080
QY 1161 TCTGGGCGCTGAAAGGGGCAAGGAGGAC 1194
Db 1081 TCTGGGCGCTGAAAGGGGCAAGGAGGAC 1114

RESULT 4
US-10-012-896-387
; Sequence 387, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepier, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-012-896-387

Query Match 33.8%; Score 522; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131; Indels 0; Gaps 0;
Matches 522; Conservative 0; Mismatches 0

QY 1025 GGCCCGAGTGGGACCAAGGAGCTTTTGCAGGCTTCTTCCAGGCTTCTTCCAGGCTG 1084
Db 1 GGCCCGAGTGGGACCAAGGAGCTTCTTCCAGGCTTCTTCCAGGCTG 60

QY 1085 CCCCCTCTGTGGCATCATGATCAGCACCTATGATTCGGCAAGGCTTCTTCCAGGCTG 1144
Db 61 CCCCCTCTGTGGCATCATGATCAGCACCTATGATTCGGCAAGGCTTCTTCCAGGCTG 120

QY 1145 TGAAACAGGACCGGCTTCTGGGCGCTGAAAGGGGCAAGGAGGCAAGGCTTCTTCC 1204
Db 121 TGAAACAGGACCGGCTTCTGGGCGCTGAAAGGGGCAAGGAGGCAAGGCTTCTTCC 180

QY 1205 CCACGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1264
Db 181 CCACGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240

QY 1265 GAGGGGCTGTGTTCCTTCCCTCCCGCGCAAGCTCCAGGCGAGGCTGTCCTCTGG 1324
Db 241 GAGGGGCTGTGTTCCTTCCCTCCCGCGCAAGCTCCAGGCGAGGCTGTCCTCTGG 300

QY 1325 GCGGCCAGGACTTCTTCCAGACACAACTTCTTCCCTGCTGCTCCAGTCTGGGATCATCA 1384
Db 301 GCGGCCAGGACTTCTTCCAGACACAACTTCTTCCCTGCTGCTCCAGTCTGGGATCATCA 360

QY 1385 CTTACCCACCCCAAGTTCAGACCAAAATCTTCCAGTCTGCCCTTCTGTTCCTCTG 1444
Db 361 CTTACCCACCCCAAGTTCAGACCAAAATCTTCCAGTCTGCCCTTCTGTTCCTCTG 420

QY 1445 GTTGTGTGTAGTGGGATGTCTTCCAGGAACCAAGAGCCCTCAGGCTGTGTGTCTCC 1504
Db 421 GTTGTGTGTAGTGGGATGTCTTCCAGGAACCAAGAGCCCTCAGGCTGTGTGTCTCC 480

QY 1505 CTCACCTTGTAAATCTTCTTAAAGTCTAAAGATGATGAATTC 1546
Db 481 CTCACCTTGTAAATCTTCTTAAAGTCTAAAGATGATGAATTC 522

RESULT 5
US-09-895-793-387
; Sequence 387, Application US/09895793
; Publication No. US20020192763A1

Db 361 CTTACCCACCCCAAGTTCAGACCAATCTTCAGCTGCCCTTCGTGTTCCCTGT 420
QY 1445 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCCCTTGTAAATCTTAAAGTCTAAAGATGATGAATTC 1546
Db 481 CTGACCCCTTGTAAATCTTAAAGTCTAAAGATGATGAATTC 522

RESULT 7

US-09-759-143-387

; Sequence 387, Application US/09759143

; Patent No. US200202248A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Ailun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C23

; CURRENT APPLICATION NUMBER: US/09759,143

; NUMBER OF SEQ ID NOS: 934

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 387

; LENGTH: 537

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-759-143-387

Query Match 33.8%; Score 522; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTCGGGACCAAGGACTTTTGCAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCGAGTCGGGACCAAGGACTTTTGCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTCCAGAGGC 1144
Db 61 CCCCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTCCAGAGGC 120
QY 1145 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 1204
Db 121 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 180
QY 1205 CCACGGATGGGAGAGGGGCAAGGAGAGCCACCAAGTGCCTTTCTCAGACTGAGG 1264
Db 181 CCACGGATGGGAGAGGGGCAAGGAGAGCCACCAAGTGCCTTTCTCAGACTGAGG 240
QY 1265 GAGGGGGTGTGTTCCCTTCCTCCCGGACAGCTCCAGGAGGCTGTCCTCTGG 1324
Db 241 GAGGGGGTGTGTTCCCTTCCTCCCGGACAGCTCCAGGAGGCTGTCCTCTGG 300
QY 1325 GGGGCCCCAGCACTTCTTCAGACCAACTTCTTCCTGTGCTCCAGTCTGTGGGATCATCA 1384

Db 301 GGGGCCGAGCACTTCCTCAGACACAACATTTCTTCCTGTGTCCAGTCTGTGGGATCATCA 360
QY 1385 CTTACCCACCCCAAGTTCAGAGACCAATCTTCAGCTGCCCTTCGTGTTCCCTGT 1444
Db 361 CTTACCCACCCCAAGTTCAGAGACCAATCTTCAGCTGCCCTTCGTGTTCCCTGT 420
QY 1445 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCCCTTGTAAATCTTAAAGTCTAAAGATGATGAATTC 1546
Db 481 CTGACCCCTTGTAAATCTTAAAGTCTAAAGATGATGAATTC 522

RESULT 8

US-09-780-669-387

; Sequence 387, Application US/09780669

; Patent No. US20020051977A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Ailun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C24

; CURRENT APPLICATION NUMBER: US/09780,669

; NUMBER OF SEQ ID NOS: 943

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 387

; LENGTH: 537

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-780-669-387

Query Match 33.8%; Score 522; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTCGGGACCAAGGACTTTTGCAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCGAGTCGGGACCAAGGACTTTTGCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTCCAGAGGC 1144
Db 61 CCCCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTCCAGAGGC 120
QY 1145 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 1204
Db 121 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 180
QY 1205 CCACGGATGGGAGAGGGGCAAGGAGAGCCACCAAGTGCCTTTCTCAGACTGAGG 1264
Db 181 CCACGGATGGGAGAGGGGCAAGGAGAGCCACCAAGTGCCTTTCTCAGACTGAGG 240

QY 1265 GAGGGGGCTGTTTCCCTCCCGGCGAAGCTCCAGGCGAGGCTGTCCCTCTGG 1324
DB 241 GAGGGGGCTGTTTCCCTCCCGGCGAAGCTCCAGGCGAGGCTGTCCCTCTGG 300
QY 1325 GGGGCCAGACATCTCTCAGACACAACTTCTTCTGCTGTCCAGTCGTGGGGATCATCA 1384
DB 301 GGGGCCAGACATCTCTCAGACACAACTTCTTCTGCTGTCCAGTCGTGGGGATCATCA 360
QY 1385 CTTTACCAGCCGCCCAAGTTCAAGACCAATCTTCCAGTGCCTCCCTTCTGTTTCCCTGT 1444
DB 361 CTTTACCAGCCGCCCAAGTTCAAGACCAATCTTCCAGTGCCTCCCTTCTGTTTCCCTGT 420
QY 1445 GTTTCCTGTAGCTGGGATGTTCTCCAGGAAACCAAGAGCCCTCAGCCTGTGTAGTCTCC 1504
DB 421 GTTTCCTGTAGCTGGGATGTTCTCCAGGAAACCAAGAGCCCTCAGCCTGTGTAGTCTCC 480
QY 1505 CTGACCCCTGTTTAACTTCTTAAGTCTAAGATGATGAATTC 1546
DB 481 CTGACCCCTGTTTAACTTCTTAAGTCTAAGATGATGAATTC 522
RESULT 9
US-09-822-827-387
; Sequence 387, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822.827
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-387
Query Match 33.8%; Score 522; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1025 GGGCGAGTGGGACCAAGAGGACTTCTTGGAGGCTTCTTCCGATCATCAAGGCTG 1084
DB 1 GGGCGAGTGGGACCAAGAGGACTTCTTGGAGGCTTCTTCCGATCATCAAGGCTG 60
QY 1085 CCCCTCTGTGCTATCATGATCAGACCTATGAGTTCCGCAAAAGCTTCTTCCAGAGGC 1144
DB 61 CCCCTCTGTGCTATCATGATCAGACCTATGAGTTCCGCAAAAGCTTCTTCCAGAGGC 120
QY 1145 TGAACCAAGACCCGCTTCTGGGCGCTGAAAGGGGCAAGGAGCAAGGACCCGCTCTC 1204
DB 121 TGAACCAAGACCCGCTTCTGGGCGCTGAAAGGGGCAAGGAGCAAGGACCCGCTCTC 180
QY 1205 CCACGATGGGAGAGGGGAGGAGGAGACCCAGCAAGTGCCTTTTCTCAGCACTGAGG 1264
DB 181 CCACGATGGGAGAGGGGAGGAGGAGACCCAGCAAGTGCCTTTTCTCAGCACTGAGG 240
QY 1265 GAGGGGGCTGTTTCCCTTCCCTCCGCGGAGACAGTCCAGGCGAGGCTGTCCCTCTGG 1324
DB 241 GAGGGGGCTGTTTCCCTTCCCTCCGCGGAGACAGTCCAGGCGAGGCTGTCCCTCTGG 300
QY 1325 GGGGCCAGACATCTCTCAGACACAACTTCTTCTGCTGTCCAGTCGTGGGGATCATCA 1384
DB 301 GGGGCCAGACATCTCTCAGACACAACTTCTTCTGCTGTCCAGTCGTGGGGATCATCA 360
QY 1385 CTTTACCAGCCGCCCAAGTTCAAGACCAATCTTCCAGTGCCTCCCTTCTGTTTCCCTGT 1444
DB 361 CTTTACCAGCCGCCCAAGTTCAAGACCAATCTTCCAGTGCCTCCCTTCTGTTTCCCTGT 420
QY 1445 GTTTCCTGTAGCTGGGATGTTCTCCAGGAAACCAAGAGCCCTCAGCCTGTGTAGTCTCC 1504

DB 421 GTTTCCTGTAGCTGGGATGTTCTCCAGGAAACCAAGAGCCCTCAGCCTGTGTAGTCTCC 480
QY 1505 CTGACCCCTGTTTAACTTCTTAAGTCTAAGATGATGAATTC 1546
DB 481 CTGACCCCTGTTTAACTTCTTAAGTCTAAGATGATGAATTC 522
RESULT 10
US-09-880-107-1133/C
; Sequence 1133, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1133
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA450247
US-09-880-107-1133
Query Match 28.8%; Score 445.4; DB 10; Length 452;
Best Local Similarity 99.8%; Pred. No. 1.8e-110;
Matches 446; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1100 TCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACACGACCCGC 1159
DB 452 TCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACACGACCCGC 393
QY 1160 TTCTGGGCGCTGAAAGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1219
DB 392 TTCTGGGCGCTGAAAGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 333
QY 1220 GGGCAGGAGGAGACCCAGCCAAAGTGCCTTTTCTCAGCACTGAGGGAGGGGCTTGTTC 1279
DB 332 GGGCAGGAGGAGACCCAGCCAAAGTGCCTTTTCTCAGCACTGAGGGAGGGGCTTGTTC 273
QY 1280 CCTTCCCTCCCGGCGACAGCTCCAGGCGAGGCTGTCCCTCTGGGGGGCCAGCACATTC 1339
DB 272 CCTTCCCTCCCGGCGACAGCTCCAGGCGAGGCTGTCCCTCTGGGGGGCCAGCACATTC 213
QY 1340 CTCAGACAACTTCTTCTGCTGTCCAGTCTGCGGGGATCATCACTTACCCACCCCCCA 1399
DB 212 CTCAGACAACTTCTTCTGCTGTCCAGTCTGCGGGGATCATCACTTACCCACCCCCCA 153
QY 1400 AGTTCAAGACCAAACTTCTCAGCTGCCCTTCTGCTGTTCCTGTGTGTGTGTGTGTGT 1459
DB 152 AGTTCAAGACCAAACTTCTCAGCTGCCCTTCTGCTGTTCCTGTGTGTGTGTGTGTGT 93
QY 1460 GCATGCTCCAGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1519
DB 92 GCATGCTCCAGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 33
QY 1520 TCCCTTAACTCTAAGATGATGAATTC 1546
DB 32 TCCCTTAACTCTAAGATGATGAATTC 6
RESULT 11

```

US-09-880-107-1179/c
; Sequence 1179, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1179
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA455522
US-09-880-107-1179
Query Match 14.1%; Score 218.4; DB 10; Length 240;
Best Local Similarity 97.1%; Pred. No. 1.6e-49;
Matches 233; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1308 CAGGGCTGTCCTCTGGGGCCGAGCACTTCTCAGACAACTTTCTCTCTCTCTCC 1367
DB 240 CAGGGCTGTCCTCTGGGGCCGAGCACTTCTCAGACAACTTTCTCTCTCTCTCC 181
QY 1368 AGTCGTGGGATCATCACTTACCCA-CCCCCAAGTTCAAGACCAAAATCTCCAGCTGCC 1426
DB 180 AGTCGTGGGATCATCACTTACCCAAGTTCAAGACCAAAATCTCCAGCTGCC 1426
QY 1427 CCCTTCGTGTTTCCCTGTTTGTGTAGCTGGGATGTTCCAGGAACCAAGACCTT 1486
DB 120 CCCTTCGTGTTTCCCTGTTTGTGTAGCTGGGATGTTCCAGGAACCAAGACCTT 61
QY 1487 CAGCCTGTGTAGTCTCCGACCCCTTGTTAATCTTAACTTAAAGATGATCACTTC 1546
DB 60 CAGCCTGTGTAGTCTCCGACCCCTTGTTAATCTTAACTTAAAGATGATCACTTC 1

RESULT 12
US-09-867-701-4486
; Sequence 4486, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Hartlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4486
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapien
; OTHER INFORMATION:
US-09-867-701-4486
Query Match 12.0%; Score 185.8; DB 10; Length 453;
Best Local Similarity 92.0%; Pred. No. 1.2e-40;
Matches 207; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 583 CCATGCTGGCTGGCGCGCTGGGACCGTCACTGTGATCAGCCCCCTGGAG 642

```

```

DB 230 CCATGCTGGCTGGCGCGCTGGGACCGTCACTGTGATCAGCCCCCTGGAG 289
QY 643 CTTATCGGACAAAGCTCAGGCTCAGCATGTGTGTACCGGAGCTGGGTGCTGTGT 702
DB 290 CTTATCGGACAAAGCTCAGGCTCAGCATGTGTGTACCGGAGCTGGGTGCTGTGT 349
QY 703 CGAACTGCAGTGGCTCAGGCTGGGCGCTCACTGTGGCTGGGCGCCCACTGCC 762
DB 350 CGAACTGCAGTGGCTCAGGCTGGGCGCTCACTGTGGCTGGGCGCCCACTGCC 408
QY 763 CTTTCGAGATGTCCTCTTCTCAGCCCTGTACTGTTCACTATGAG 807
DB 409 CTTTCGAGATGTCCTCTTCTCAGCCCTGTACTGTTCACTATGAG 453

RESULT 13
US-09-908-711-46
; Sequence 46, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA128
; CURRENT APPLICATION NUMBER: US/09/908,711
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US01/01360
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,867
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01344
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01345
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,905
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01354
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01339
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01340
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,874
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01334
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,898
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01320
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,853
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01349
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,902
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01239
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01348
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,882

```



```
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 11-LIB34-070-Q1-E1-C11
US-09-960-352-2543

Query Match      10.4%; Score 161; DB 10; Length 246;
Best Local Similarity 87.8%; Pred. No. 4,4e-34;
Matches 187; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

Qy  49 ATGGGCGCGCTGGGCGCTGGGCGCGCGCGCACGAGCACCGCTAGAGCCAGGACTGA 108
Db  9  ATGGGTCCGCGCGCGCGCGCGCGCGCGCTCGAGCACCGGCGCTAGAGCCAGAACTGA 68

Qy  109 AGCTTCAGATGGCTGACCGAGGACCTGGGGCATCAGCCCCCTCCAGCAATGGTGGCC 168
Db  69 AGCTACAGATGGCTGACCGAGGACCCAGGGGGCATCAGCCCTCTCCAAACAATGGTGGCA 128

Qy  169 TCAGGCACCGGGGCTGTGGTTACCTCTCTTCATGACACCCCTGGACGTGGTGAAGGTT 228
Db  129 TCCGGCGCGGGGCTGTGTACCTCTCTTCATGACACCCCTGGATCTGGTGAAGGTA 188

Qy  229 CGCTTGCAG-TCTCAGCGGCGCTCCATGGCCAG 260
Db  189 CGTTTGCAGTTCTCAGCGCGCGCTCAGTGGCCAG 221
```

Search completed: January 22, 2003, 19:08:08
Job time : 91.3 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 19:06:17 ; Search time 2432.33 Seconds
(without alignments)
10293.906 Million cell updates/sec

Title: US-09-888-358-2

Perfect score: 1546

Sequence: 1 Ggctaggctgcgtgcgagcgcg.....gtctaaagatgaacttc 1546

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Tc number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_esti:*
9: gb_esti:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1381.2	89.3	1402	11 AF119864	AF119864 Homo sapi
2	909.6	58.8	1020	13 BM554056	BM554056 AGENCOURT
3	908.8	58.8	963	9 AL519986	AL519986 AL519986
4	888.8	57.5	981	14 BM923995	BM923995 AGENCOURT
5	878.8	56.8	1001	9 AU519987	AL519987 AU519987
6	877.4	56.8	921	9 AU529906	AL529906 AU529906

7	869.4	56.2	920	14 BQ683596	BQ683596 AGENCOURT
8	869	56.2	932	14 BQ925174	BQ925174 AGENCOURT
9	868.4	56.2	925	9 AL516722	AL516722 AL516722
10	867	56.1	911	9 AL528266	AL528266 AL528266
11	864	55.9	910	9 AL570513	AL570513 AL570513
12	855.8	55.4	959	14 BM915493	BM915493 AGENCOURT
13	854	55.2	939	14 BQ949458	BQ949458 AGENCOURT
14	853.8	55.2	965	14 BQ934841	BQ934841 AGENCOURT
15	847	54.8	885	9 AL514846	AL514846 AL514846
16	839.6	54.3	1077	13 BM808978	BM808978 AGENCOURT
17	838.8	54.3	969	14 BI759919	BI759919 603044343
18	837.2	54.2	884	9 AL572559	AL572559 AL572559
19	836.6	54.1	901	9 AL545150	AL545150 AL545150
20	828	53.6	1120	13 BM464148	BM464148 AGENCOURT
21	826.2	53.4	1450	11 AK007934	AK007934 Mus muscu
22	823	53.2	1080	14 BM917004	BM917004 AGENCOURT
23	820.4	53.1	891	9 AL522522	AL522522 AL522522
24	820.2	53.1	890	9 AL515130	AL515130 AL515130
25	818.6	52.9	994	13 BM555499	BM555499 AGENCOURT
26	818	52.9	1012	13 BM468665	BM468665 AGENCOURT
27	813.4	52.6	962	14 BQ688847	BQ688847 AGENCOURT
28	813.4	52.6	1112	14 BM810266	BM810266 AGENCOURT
29	812	52.5	917	9 AL544360	AL544360 AL544360
30	811.4	52.5	1371	11 AK007276	AK007276 Mus muscu
31	809	52.3	872	14 BQ883416	BQ883416 AGENCOURT
32	805.8	52.1	1103	14 BM924115	BM924115 AGENCOURT
33	805	52.1	911	12 BG337246	BG337246 60243762
34	802	51.9	816	12 BG470240	BG470240 602533758
35	801.6	51.8	1453	11 AK002381	AK002381 Mus muscu
36	797.2	51.6	1004	14 BM928609	BM928609 AGENCOURT
37	795.8	51.5	940	9 AL523298	AL523298 AL523298
38	790.2	51.1	847	13 BI333101	BI333101 602996377
39	789.4	51.1	862	12 BG470248	BG470248 602533766
40	788.8	51.0	829	9 AL563104	AL563104 AL563104
41	788.8	51.0	1092	13 BM460198	BM460198 AGENCOURT
42	787.2	50.9	1491	11 AK019396	AK019396 Mus muscu
43	787	50.9	1029	13 BM459839	BM459839 AGENCOURT
44	786.4	50.9	918	12 BG574617	BG574617 602596614
45	785	50.8	795	9 AL571090	AL571090 AL571090

ALIGNMENTS

RESULT 1
AF119864
LOCUS AF119864 1402 bp mRNA linear HTC 08-MAY-2001
DEFINITION Homo sapiens PRO2163 mRNA, complete cds.
ACCESSION AF119864
VERSION AF119864.1 GI:7770164
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
and He,F.
TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1402)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
and He,F.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing, Beijing 100850, P.R. China
FEATURES
Location/Qualifiers
1..1402
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLB8041"

CDS	/tissue_type="liver"	
	/dev_stage="fetus"	
	4..1017	
	/note="predicted protein of HQ2163"	
	/codon_start=1	
	/product="PRO2163"	
	/protein_id="AAF69618.1"	
	/db_xref="GI:7770165"	
	/translation="MVASGTGAVVTSLFMTPLDVVKVRLQSRQSMASELMPSSRLWS LSYTKWKLCLYCNGLVLEPLFLYPCNGARCATFODPRTFTGMDAFVKIVRHEGTRTLW SLPATLVMTVPATAIYFYADQLKAFELCFLMSLTDLXAPMVAGALARKIVTVISPL ELMRKLOHVSYSRELGACVRYTAVAGGWRSLMLGWGPTALRDVPSALYWNVELV KSWNLGRPKDQTSVGMFSFVAGG:SRVRAVLTLPEDEVKTOBOVALGMEAVRNPL HVSTWLLLRIRAESGKGLPAGFLPRIIKAAPSCAIMISTYEFKSGFFQRLNQDRL LOG"	
BASE COUNT	261 a	432 c 401 g 308 t
ORIGIN		
Query Match	89.3%;	Score 1381.2; DB 11; Length 1402;
Best Local Similarity	99.8%;	Pred. No. 0;
Matches 1383; Conservative	0;	Mismatches 3; Indels 0; Gaps 0;
QY	157	CAAAATGGTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTTCATGACACCCCTGGAC 216
Db	1	CAAAATGGTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGAC 60
QY	217	GTGGTGAAGTTGGCTGCGAGTCTCAGCGGCGCTCCATGGCCAGCAGCTGATGCTCTCC 276
Db	61	GTGGTGAAGTTGGCTGCGAGTCTCAGCGGCGCTCCATGGCCAGCAGCTGATGCTCTC 120
QY	277	TCCAGACTGTGAGGCTCTCTTATACAAATGGAAGTGCCTCTGTATATGCAATGGTGTC 336
Db	121	TCCAGACTGTGAGGCTCTCTTATACAAATGGAAGTGCCTCTGTATATGCAATGGTGTC 180
QY	337	CTGGAGGCTCTGTACCTGTGCCAAATGGTGCCCGCTGTGCCACCTGGTTTCAAGACCT 396
Db	181	CTGGAGGCTCTGTACCTGTGCCAAATGGTGCCCGCTGTGCCACCTGGTTTCAAGACCT 240
QY	397	ACCCGCTTCACTGGCACCATGGATGCCCTTCTGTGAAGATCGTGAGCGACAGGCGACGAG 456
Db	241	ACCCGCTTCACTGGCACCATGGATGCCCTTCTGTGAAGATCGTGAGCGACAGGCGACGAG 300
QY	457	ACCTCTTGAGCGGCGCTCCCGCCACCTGGTGATGACTGTGCCAGTACCGGCATCTTAC 516
Db	301	ACCTCTTGAGCGGCGCTCCCGCCACCTGGTGATGACTGTGCCAGTACCGGCATCTTAC 360
QY	517	TTCACTGCTTATGACCAACTGAAGGCTTCTGTGTGGTTCAGGACCTGACTCTGACCTC 576
Db	361	TTCACTGCTTATGACCAACTGAAGGCTTCTGTGTGGTTCAGGACCTGACTCTGACCTC 420
QY	577	TACGACCCCACTGGTGGCTGGCGGCTGGCCCGCTGGGCACCGTGACTGTGATCAGCCCC 636
Db	421	TACGACCCCACTGGTGGCTGGCGGCTGGCCCGCTGGGCACCGTGACTGTGATCAGCCCC 480
QY	637	CTGGAGCTTATGCGGCAAAAGCTGCAGGCTCAGCATGTGTCTGTACCGGAGCTGG-ITGCC 696
Db	481	CTGGAGCTTATGCGGCAAAAGCTGCAGGCTCAGCATGTGTCTGTACCGGAGCTGGTGGCT 540
QY	697	TGTGTTTGAACCTGCACTGGCTCAGGGTGGCTGGCGCTCACTGTGGCTGGGCTGGGCCCC 756
Db	541	TGTGTTTGAACCTGCACTGGCTCAGGGTGGCTGGCGCTCACTGTGGCTGGGCTGGGCCCC 600
QY	757	ACTGCCCCTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGCTGGTGAAG 816
Db	601	ACTGCCCCTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGCTGGTGAAG 660
QY	817	AGCTGGCTCAATGGGTTTCAGGCCGAAGACACAGACTTCTCTGGGCATGAGCTTGTGGCT 876
Db	661	AGCTGGCTCAATGGGCTCAGGCCGAAGACACAGACTTCTCTGGGCATGAGCTTGTGGCT 720
QY	877	GTTGGCATCTCAGGACCGGTGCTGCACTGCTGACTCTACCTTTTGACGTGGTAAAGACC 936
Db	721	GTTGGCATCTCAGGACCGGTGCTGCACTGCTGACTCTACCTTTTGACGTGGTAAAGACC 780

Qy	937	CAACGCCAGTGGCTCTGGAGCGATGGAGGCTGTGAGAGTGAACCCCTGCAATGTGGAC	996
Db	781	CAACGCCAGTGGCTCTGGAGCGATGGAGGCTGTGAGAGTGAACCCCTGCAATGTGGAC	840
Qy	997	TCCACCTGGCTGCTGCGGAGGATCCGGGCGGAGTGGGCGACCAAGGACTCTTTGCA	1056
Db	841	TCCACCTGGCTGCTGCGGAGGATCCGGGCGGAGTGGGCGACCAAGGACTCTTTGCA	900
Qy	1057	GGCTTCCTTCCTCGGATCATCAAGGCTGCCCTCTCTGCGCATCATGATCAGCACTAT	1116
Db	901	GGCTTCCTTCCTCGGATCATCAAGGCTGCCCTCTCTGCGCATCATGATCAGCACTAT	960
Qy	1117	GAGTTCGGCAAAAGCTTCTTCAGAGGCTGAAACAGGACCGGCTTCTGGCGGCTGAAAG	1176
Db	961	GAGTTCGGCAAAAGCTTCTTCAGAGGCTGAAACAGGACCGGCTTCTGGCGGCTGAAAG	1020
Qy	1177	GGCAAGGAGGACGAGGACCGCTCTCTCCACGATGGGAGAGGCGAGGAGACCA	1236
Db	1021	GGCAAGGAGGACGAGGACCGCTCTCTCCACGATGGGAGAGGCGAGGAGACCA	1080
Qy	1237	GCCAGTGCCTTTTCTCAGCACTGAGGAGGGGCTTGTTCCTTCCCTCCCGGCGAC	1296
Db	1081	GCCAGTGCCTTTTCTCAGCACTGAGGAGGGGCTTGTTCCTTCCCTCCCGGCGAC	1140
Qy	1297	AAGTTCAGGAGGAGGCTGCTCTCTGGGCGGCGCAGCACTTCTCAGACAACTTCTT	1356
Db	1141	AAGTTCAGGAGGAGGCTGCTCTCTGGGCGGCGCAGCACTTCTCAGACAACTTCTT	1200
Qy	1357	CTGTGCTCCAGTCTGGGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCT	1416
Db	1201	CTGTGCTCCAGTCTGGGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCT	1260
Qy	1417	TCCAGTGCCTTCGCTGCTTCCCTGCTGCTTGTGCTAGCTGGGCACTCTCCAGGAAC	1476
Db	1261	TCCAGTGCCTTCGCTGCTTCCCTGCTGCTTGTGCTAGCTGGGCACTCTCCAGGAAC	1320
Qy	1477	AAGAAGCTCAGCTGGTGTAGTCTCCCTGAGCTTGTAACTTCTTAAGTCAAGAT	1536
Db	1321	AAGAAGCTCAGCTGGTGTAGTCTCCCTGAGCTTGTAACTTCTTAAGTCAAGAT	1380
Qy	1537	GATGAA 1542	
Db	1381	GATGAA 1386	
RESULT 2			
BM554056			
LOCUS			
DEFINITION			
5', mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
<p>BM554056 1020 bp mRNA linear EST 20-FEB-2003</p> <p>AGENCOURT_6546709 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742317</p> <p>5', mRNA sequence.</p> <p>BM554056</p> <p>BM554056.1 GI:18793328</p> <p>EST.</p> <p>human.</p> <p>Homo sapiens</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>1 (bases 1 to 1020)</p> <p>NIH-MGC http://mgi.nci.nih.gov/.</p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished (1999)</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: cgapss-x@mail.nih.gov</p> <p>Tissue Procurement: Life Technologies, Inc.</p> <p>cDNA Library Preparation: Life Technologies, Inc.</p> <p>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Agencourt Bioscience Corporation</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov</p> <p>Plate: LLAM12760 row: e column: 06</p> <p>High quality sequence stop: 691.</p>			

QY 623 CTGTGATCAGCCCTCGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTCTTACC 682
 Db 904 CTGTGATCAGCCCTCGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTCTTACC 845
 QY 683 GGGAGCTGGTGGCTGTGTTCGAACATGCAAGTGGCTCAGGCTGGCTGGCTCAGTGTGC 742
 Db 844 GGGAGCTGGTGGCTGTGTTCGAACATGCAAGTGGCTCAGGCTGGCTGGCTCAGTGTGC 785
 QY 743 TGGGCTGGGGCCCACTGCTCCTTCGAGATGTGCCCTTTCAGCCCTGTACTGTTCAACT 802
 Db 784 TGGGCTGGGGCCCACTGCTCCTTCGAGATGTGCCCTTTCAGCCCTGTACTGTTCAACT 726
 QY 803 ATGAGCTGGTGAAGCTGGCTCAATGGCTTCAGGCCGAGGACAGACTTCTGTGGCA 862
 Db 725 ATGAGCTGGTGAAGCTGGCTCAATGGCTTCAGGCCGAGGACAGACTTCTGTGGCA 666
 QY 863 TGAGCTTTGTGGCTGGTGGCTCTCAGGACGGTGGCTGCAAGTGGCTGACTTCACTTACCCTTTG 922
 Db 665 TGAGCTTTGTGGCTGGTGGCTCTCAGGACGGTGGCTGCAAGTGGCTGACTTCACTTACCCTTTG 606
 QY 923 ACGTGTAAAGCAACCAAGCAAGTGGCTTCGAGGAGATGGAGCTGTGAGAGTGAACC 982
 Db 605 ACGTGTAAAGCAACCAAGCAAGTGGCTTCGAGGAGATGGAGCTGTGAGAGTGAACC 546
 QY 983 CCTGCTGATGAGCTCCACCTGGCTGCTGCGGAGGATCCGGCCGAGTGGGACCA 1042
 Db 545 CCTGCTGATGAGCTCCACCTGGCTGCTGCGGAGGATCCGGCCGAGTGGGACCA 486
 QY 1043 AGGAGCTTTTCAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTGCAATCA 1102
 Db 485 AGGAGCTTTTCAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTGCAATCA 426
 QY 1103 TGATCAGCACTATGAGTTCGCAAAAGCTTTTCCAGAGCTGAACAGGACCGGCTTC 1162
 Db 425 TGATCAGCACTATGAGTTCGCAAAAGCTTTTCCAGAGCTGAACAGGACCGGCTTC 366
 QY 1163 TGGGGCTGAAAGGGCAAGGAGGCAAGGACCCGCTCTCCACGATGGGGAGGG 1222
 Db 365 TGGGGCTGAAAGGGCAAGGAGGCAAGGACCCGCTCTCCACGATGGGGAGGG 306
 QY 1223 CAGGAGGAGACCCAGCAAGTGTCTTCTCAGCACTGAGGAGGGGCTGTTCCT 1282
 Db 305 CAGGAGGAGACCCAGCAAGTGTCTTCTCAGCACTGAGGAGGGGCTGTTCCT 246
 QY 1283 TCCTCCCGGGCAAGCTCCAGGAGGCTGTCTCTGCGGGCCCGAGCACTTCCTC 1342
 Db 245 TCCTCCCGGGCAAGCTCCAGGAGGCTGTCTCTGCGGGCCCGAGCACTTCCTC 186
 QY 1343 AGACAACTTCTTCTGCTTCCAGTGTGGGATCATCACTTACCCACCCCAAGT 1402
 Db 185 AGACAACTTCTTCTGCTTCCAGTGTGGGATCATCACTTACCCACCCCAAGT 126
 QY 1403 TCAAGCAAAATCTTCAGGTCGCCCTTCTGTTTCCTGTGTTGCTGTAGTGGCA 1462
 Db 125 TCAAGCAAAATCTTCAGGTCGCCCTTCTGTTTCCTGTGTTGCTGTAGTGGCA 67
 QY 1463 TGCTCCAGGAACCAAGAGGCTCAGGCTGGTGTAGTCTCCCTGACCTTGTATTC 1522
 Db 66 TGCTCCAGGAACCAAGAGGCTCAGGCTGGTGTAGTCTCCCTGACCTTGTATTC 7
 QY 1523 TTAA 1526
 Db 6 TAAD 3

RESULT 4
 BM923995
 LOCUS BM923995 981 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGNCOURT_6709842 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760154
 5', mRNA sequence.
 ACCESSION BM923995
 VERSION BM923995.1 GI:19374374

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 981)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-x@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12806 Row: 1 column: 11
 High quality sequence stop: 701.
 FEATURES
 Location/Qualifiers
 1..981
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5760154"
 /clone_lib="NIH_MGC_116"
 /lab_host="DH10B"
 /notes="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 cligo-df primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen) Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."
 BASE COUNT 171 a 309 c 296 g 201 t 2 others
 ORIGIN

Query Match 57.3%; Score 888.8; DB 14; Length 981;
 Best Local Similarity 97.2%; Pred. No. 1.1e-207;
 Matches 946; Conservative 0; Mismatches 23; Indels 4; Gaps 4;
 QY 33 AGGGCGGAGCAGTATGGGCGCGCTGGCCCTGGGCGCGCGCCGACGACACACG 92
 Db 5 AGGGCGGAGCAGTATGGGCGCGCTGGCCCTGGGCGCGCGCCGACGACACACG 64
 QY 93 CTAGAGCCAGGACTGAAGCTTCAAGATGCTCAGCAGGACCTCGGGGATCAGCCCT 152
 Db 65 CTAGAGCAGGACTGAAGCTTCAAGATGCTCAGCAGGACCTCGGGGATCAGCCCT 124
 QY 153 CCAGCAAAATGTTGGCCTCAGGACCGGGGCTGTGGTTACCTTCTCTTCAATGACCCCT 212
 Db 125 CCAGCAAAATGTTGGCCTCAGGACCGGGGCTGTGGTTACCTTCTCTTCAATGACCCCT 184
 QY 213 GGAGCTGGTGAAGTTGCGCTCAGTCTCAGGCGCCCTCAGTCCAGCAGCAGTATGCC 272
 Db 185 GGAGCTGGTGAAGTTGCGCTCAGTCTCAGGCGCCCTCAGTCCAGCAGCAGTATGCC 244
 QY 273 TTCTCCAGACTTGGAGCCTCTCTATACCAATGGAAGTGCCTCTGTATTGCAATGG 332
 Db 245 TTCTCCAGACTTGGAGCCTCTCTATACCAATGGAAGTGCCTCTGTATTGCAATGG 304
 QY 333 TGTCTGGAGCCTGTGACTGTGCCCCAAATGGTCCCGCTGTGCCACCTGGTTTCAAGA 392
 Db 305 TGTCTGGAGCCTGTGACTGTGCCCCAAATGGTCCCGCTGTGCCACCTGGTTTCAAGA 364
 QY 393 CCTACCCGCTTCACTGGCACCATTGATGCTTCTGTAAGATCTGAGGACGAGGGCAC 452
 Db 365 CCTACCCGCTTCACTGGCACCATTGATGCTTCTGTAAGATCTGAGGACGAGGGCAC 424

```

Qy 453 CAGGACCTCTGGAGCGGCTCCGCCACCTGGTGTGATGATGTGCGAGCTACCGCAT 512
Db 425 CAGGACCTCTGGAGCGGCTCCGCCACCTGGTGTGATGATGTGCGAGCTACCGCAT 484
Qy 513 CTACTTCACTGCTATGACCAACTGAAGCCTTCTGTGTGCTGAGCCTGACCTTGA 572
Db 485 CTACTTCACTGCTATGACCAACTGAAGCCTTCTGTGTGCTGAGCCTGACCTTGA 544
Qy 573 CTTCTACGACCCATGTGTGGCTGGCGCTGGCCGCTGGCCGCTGGGACCTGATGTATG 632
Db 545 CTTCTACGACCCATGTGTGGCTGGCGCTGGCCGCTGGCCGCTGGGACCTGATGTATG 604
Qy 633 CCCCCTGGAGCTTATGCGGCAAAAGCTGCAGGCTCAGCATGTGTGTAACCGGAGCTGG 692
Db 605 CCCCCTGGAGCTTATGCGGCAAAAGCTGCAGGCTCAGCATGTGTGTAACCGGAGCTGG 664
Qy 693 TGCCTGTGTTCAACTGCACTGAGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCT 752
Db 665 TGCCTGTGTTCAACTGCACTGAGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCT 724
Qy 753 CCCCCTGCTGCTGAGATGTGCCCTTCTCAGCCTGTACTGTTCAACTATGAGCTGGT 812
Db 725 CCCCCTGCTGCTGAGATGTGCCCTTCTCAGCCTGTACTGTTCAACTATGAGCTGGT 784
Qy 813 GAAGAGCTGGTCAATGGGTTCAAGCGG-C-AAGGACCAAGCTTCTGTGGGCAATGAGC-TTT 870
Db 785 GAAGAGCTGGTCAATGGGTTCAAGCGG-GAAGGACCAAGCTTCTGTGGGCAATGAGC-TTT 844
Qy 871 GTGGCTGTGGGATCTCAGGAGCGGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 929
Db 845 GTGGCTGTGGGATCTCAGGAGCGGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 904
Qy 930 AAGAGCCCAAGCCAGCTGCTCTGGGAGGAT-CGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 988
Db 905 AAGAGCCCAAGCCAGCTGCTCTGGGAGGATCGGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 964
Qy 989 ATGTGAGTCCAC 1001
Db 965 ATGGGAGTCCAC 977

RESULT 5
AL519987 1001 bp mRNA linear EST 13-FEB-2001
LOCUS AL519987 LTI_NFL004_NEC2 Homo sapiens cDNA clone CS0DB005YP09 5
DEFINITION prime, mRNA sequence.
ACCESSION AL519987
VERSION AL519987.1 GI:12783480
X RDS EST.
S _JE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1001)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB005YP09"
/tissue="male"
/lab_host="PH108"
/lab_host="neuroblastoma cells"
/notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and

```

```

BASE COUNT 167 a 309 c 315 g 208 t 2 others
ORIGIN
Query Match 56.8%; Score 878.8; DB 9; Length 1001;
Best Local Similarity 97.6%; Pred. No. 3.3e-205;
Matches 955; Conservative 0; Mismatches 17; Indels 6; Gaps 6;
Qy 28 CCACGAGGCGGACGGACGTAATGGGCGGCGCTGGCCCTGGGCGCGCGCCGACGACGCA 87
Db 11 CCACGAGGCGGCGGASGTTATTGGGCGGCGCTGGCCCTGGGCGCGCGCCGACGACGCA 70
Qy 88 CCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGTGTGACACGAGACCTGCGGGCATCAGC 147
Db 71 CCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGTGTGACACGAGACCTGCGGGCATCAGC 130
Qy 148 CCCTCCAGCAAAATGTGGGCTCAGGACCGGGGTGTGGTTTACCTCTCTCTTCATGACA 207
Db 131 CCCTCCAGCAAAATGTGGGCTCAGGACCGGGGTGTGGTTTACCTCTCTCTTCATGACA 190
Qy 208 CCCTGGAGGTGTGAAGTTCGCTTCAAGTGTGAGTGTGAGGCGGCTTCAATGGCCAGGAGCTG 267
Db 191 CCCTGGAGGTGTGAAGTTCGCTTCAAGTGTGAGTGTGAGGCGGCTTCAATGGCCAGGAGCTG 250
Qy 268 ATGCTTCTCCAGACTGTGGAGCTCTCTATACAAATGGAAGTGCCTCCTGTATTTC 327
Db 251 ATGCTTCTCCAGACTGTGGAGCTCTCTATACAAATGGAAGTGCCTCCTGTATTTC 310
Qy 328 AATGTGTCTGGAGGCTGTGTACCTGTGCCCCAAATGTGTGCGGCTGTGCGACCTGGTTT 387
Db 311 AATGTGTCTGGAGGCTGTGTACCTGTGCCCCAAATGTGTGCGGCTGTGCGACCTGGTTT 370
Qy 388 CAAGACCTTACCGGTTCACTGGGACCATGGATGCTTCTGTGAAGATCGTGAGGACGAG 447
Db 371 CAAGACCTTACCGGTTCACTGGGACCATGGATGCTTCTGTGAAGATCGTGAGGACGAG 430
Qy 448 GGACCCAGGACCTCTGGAGCGGCTCCCGCACCTGGTGTGATGCTGTGCCAGTACC 507
Db 431 GGACCCAGGACCTCTGGAGCGGCTCCCGCACCTGGTGTGATGCTGTGCCAGTACC 490
Qy 508 GCCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTGAGGCGCTGACC 567
Db 491 GCCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTGAGGCGCTGACC 550
Qy 568 TCTGACCTTACGACCCATGTGGTGGGCGGCTGGCCGCTGGGACCGCTGTGACTGTG 627
Db 551 TCTGACCTTACGACCCATGTGGTGGGCGGCTGGCCGCTGGGACCGCTGTGACTGTG 610
Qy 628 ATCAGGCGGCTGGAGCTTATGCGGCAAAAGCTGCAGGCTCAGCATGTGTGATCCGGGAG 687
Db 611 ATCAG-CCCTGGAGCTTATGCGGCAAAAGCTGCAGGCTCAGCATGTGTGATCCGGGAG 669
Qy 688 CTGGGTGCTGTGTGTGAACTGCAAGTGGCTCAGGCTGGCTGGGCTCACTGTGGCTGGGC 747
Db 670 CTGGGTGCTGTGTGTGAACTGCAAGTGGCTCAGGCTGGCTGGGCTCACTGTGGCTGGGC 729
Qy 748 TGGGGCCCACTGCGCTTCAAGATGTGCCCTTCTCAGCCCTGTACTGGTCAACTATGAG 807
Db 730 TGGGG-CCCACTGCGCTTCAAGATGTGCCCTTCTCAGCCCTGTACTGGTCAACTATGAG 788
Qy 808 CTGCTGAAGAGCTGGCTCAATGGGTTCAAGGACGAGCTTCTGTGGGCAATGAGC 867
Db 789 CTGCTGAAGAGCTGGCTCAATGGGTTCAAGGACGAG- GACCAAGCTTCTGTGGGCAATGAGC 847
Qy 968 TTTGTGGCTGTGGCATCTCAGGAGCGGTGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 927
Db 848 TTTGTGG- TGTGTGTCATCTCAGGAGCGGTGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 907

```

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

```

QY 928 GTAAAGACCCAAACCCAGGTCGCTCTGGAGCGATGAGAGCTGTGAGAGTGAAACCCCTG 987
    |||||
Db 908 GTAAAGA-CAACCCAGGTCGCTCTGGAG-GATGGGAGGTGTGAGAGTGAA-CCCCTG 964
    |||||
QY 988 CATGTGGACTCCACTGG 1005
    |||||
Db 965 CATGTGGACTCCACTGG 982
    |||||

RESULT 6
AL529906 921 bp mRNA linear EST 13-FEB-2001
LOCUS AL529906 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD005YL19 5
DEFINITION prime, mRNA sequence.
ACCESSION AL529906
VERSION AL529906.1 GI:12793399
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 921)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
    1..921
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CS0DD005YL19"
        /clone_lib="LTI_NFL001_NBC4"
        /sex="male"
        /tissue_type="neuroblastoma cells"
        /lab_host="DH10B"
        /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
        was primed with a NotI-oligo(dT) primer. Five prime end
        enriched, double-stranded cDNA was digested with Not I and
        cloned into the Not I and Eco RV sites of the pCMVSPORT 6
        vector. Library was normalized. Library was constructed
        by Life Technologies. Contact : Feng Liang Life
        Technologies, a division of Invitrogen 9800 Medical Center
        Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
        8371 Email : fliang@lifetech.com URL :
        http://fulllength.invitrogen.com"
BASE COUNT 154 a 292 c 269 g 193 t 13 others
ORIGIN
    Query Match 56.8%; Score 877.4; DB 9; Length 921;
    Best Local Similarity 97.5%; Pred. No. 7e-205;
    Matches 889; Conservative 10; Mismatches 12; Indels 1; Gaps 1;

QY 35 GCGGACGACGCTAATGGCCCGCCCTGGCCCTGGCCGCGCGGCGGACGACGACGCT 94
    |||||
Db 11 GCGGCGGASGANTATTGGCCCGCCCTGGCCCTGGCCGCGCGGCGGACGACGACGCT 70
    |||||
QY 95 AGAGCAGGACTGAAGCTTCAAGATGGCTGACGAGCCCTGGGGCATCAGCCCTCC 154
    |||||
Db 71 AGAGCCAGGACTGAAGCTTCAAGATGGCTGACGAGCCCTGGGGCATCAGCCCTCC 130
    |||||
QY 155 AGCAATGGTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTCTTTCATCACACCCCTGG 214
    |||||
Db 131 AGCAATGGTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTCTTTCATCACACCCCTGG 190
    |||||
QY 215 ACGTGGTGAAGTTGGCTGAGTCTCAGGCGCCCTCATCGGCGAGGAGCTGATGCCTT 274
    |||||
Db 191 ACGTGGTGAAGTTGGCTGAGTCTCAGGCGCCCTCATCGGCGAGGAGCTGATGCCTT 250
    |||||
QY 275 CCTCCAGACTGTGGAGCTCTCTATACCAATGGAAGTCTCTCTATGCAATGGTG 334
    |||||

```

```

Db 251 CCTCAGACTGTGAGGCTCTCTATACCAATGGAAGTGCCTCTGTATTCGAATGGTG 310
    |||||
QY 335 TCTTGAGACCTCTGTACTGTGCCCCAAATGGTGGCCGCTGTGCCACCTGTGTTCAAGACC 394
    |||||
Db 311 TYCTGAGGCTCTGTACTGTGCCCCAAATKGTGGCCGCTGTGCCACCTGTGTTCAAGACC 370
    |||||
QY 395 CTACCCGTTTCACTGGCACCATGATGCTTCTGTTGAAGATCGTGAGGACGAGGACCA 454
    |||||
Db 371 CTA/CCGTTTCACTGGCACCATGATGCTTCTGTTGAAGATCGTGAGGACGAGGACCA 430
    |||||
QY 455 GGACCTCTGAGGCGCTCCCGCACCCCTGTTGATGACTGTGCCAGCTACCGCATCT 514
    |||||
Db 431 GGACCTCTGAGGCGCTCCCGCACCCCTGTTGATGACTGTGCCAGCTACCGCATCT 490
    |||||
QY 515 ACTTCACCTGCTATGACCAACTGAAGGCTTCTCTGTTGTTGAGGCGCTGACCTCTGACC 574
    |||||
Db 491 ACTTCACCTGCTATGACCAACTGAAGGCTTCTCTGTTGTTGAGGCGCTGACCTCTGACC 550
    |||||
QY 575 TCTAGCACCATGTTGGCTGGCGCGCTGGCCGCTGGGACCGTGACTGTGATCAGCC 634
    |||||
Db 551 TCTAGCACCATGTTGGCTGGCGCGCTGGCCGCTGGGACCGTGACTGTGATCAGCC 610
    |||||
QY 635 CCTCGAGCTTATCGGACAAAGCTGACAGGCTCAGCATGTGTTACCGGAGCTGGTG 694
    |||||
Db 611 CCTCGAGCTTATCGGACAAAGCTGACAGGCTCAGCATGTGTTACCGGAGCTGGTG 670
    |||||
QY 695 CTGTGTTTGAACCTGACAGTGGCTCAGGTTGGCTGGCGCTCACTGTGGCTGGGCTGGGCC 754
    |||||
Db 671 CTGTGTTTGAACCTGACAGTGGCTCAGGTTGGCTGGCGCTCACTGTGGCTGGGCTGGGCC 730
    |||||
QY 755 CCACCTGCTTCAGATGTGCTCTCAGCCCTGACTGTTCAACTATGAGCTGGTGA 814
    |||||
Db 731 CCACCTGCTTCAGATGTGCTCTCAGCCCTGACTGTTCAACTATGAGCTGGTGA 790
    |||||
QY 815 AGAGTGGCTCAATGGGTTTCAGGCGCGGACGACACTTCTGTGGGCGATGAGCTTGTGG 874
    |||||
Db 791 AGAGTGGCTCAATGGGTTTCAGGCGCGGACGACACTTCTGTGGGCGATGAGCTTGTGG 849
    |||||
QY 875 CTGTGGGCTCTCAGGACCGTGGCTGACGCTGCTACCTTACCTTGTGAGCTGGTAAAG 934
    |||||
Db 850 CTGTGGGCTCTCAGGACCGTGGCTGACGCTGCTACCTTACCTTGTGAGCTGGTAAAG 909
    |||||
QY 935 CCCAAGCGCAGG 946
    |||||
Db 910 CCCAAGCGCAGG 921
    |||||

RESULT 7
BO683596 920 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8059897 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091198
DEFINITION 5', mRNA sequence.
ACCESSION BO683596
VERSION BO683596.1 GI:21796275
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium information can be
http://image.llnl.gov

```


Plate: LLCM2333 row: m column: 23
High quality sequence stop: 697.

FEATURES
source
1..920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6091198"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 161 a 288 c 273 g 196 t 2 others
OF: 'N

Query Match 56.2%; Score 869.4; DB 14; Length 920;
Best Local Similarity 99.7%; Pred. No. 6.5e-203;
Matches 881; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 75 CGCCGACGAGCACCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACGAGGACCC 134
Db 1 CGCCGACAGCAGCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACGAGGACCC 60

QY 135 TCGGGGATGACGCCCTCCAGCAATGTGGCTCAGGACCGGGCTGTGGTTACCTC 194
Db 61 TCGGGGATGACGCCCTCCAGCAATGTGGCTCAGGACCGGGCTGTGGTTACCTC 120

QY 195 TCTCTTCATGACACCCCTGAGCTGTGAAGTTTCGCTCGAGTCTCAGCGCCCTCCAT 254
Db 121 TCTCTTCATGACACCCCTGAGCTGTGAAGTTTCGCTCGAGTCTCAGCGCCCTCCAT 180

QY 255 GSCCAGGAGCTGATGCTTCTCCAGACTGTGGAGCTTCTCTATACCAATGGAAGTG 314
Db 181 GSCCAGGAGCTGATGCTTCTCCAGACTGTGGAGCTTCTCTATACCAATGGAAGTG 240

QY 315 CTCTCTGATTTGCAATGGTCTCTGAGCTCTGACCTCTGACCTGTGCCAAATGGTGGCGGTG 374
Db 241 CTCTCTGATTTGCAATGGTCTCTGAGCTCTGACCTCTGACCTGTGCCAAATGGTGGCGGTG 300

QY 375 TGCCACCTGGTTTCAAGACCTTACCGCTTCACTGGACACCATGATGCTCTGTGAAGAT 434
Db 301 TGCCACCTGGTTTCAAGACCTTACCGCTTCACTGGACACCATGATGCTCTGTGAAGAT 360

QY 435 CQTGAGGACAGAGGACACGACCTCTGAGCGGCTTCCCGCCACCCCTGTGTGATGAC 494
Db 361 CQTGAGGACAGGAGGACACGACCTCTGAGCGGCTTCCCGCCACCCCTGTGTGATGAC 420

QY 495 TGTGACGATCAGCGATCTTACTTCTGCTATGACCACTGAGGCTTCTGTGTGG 554
Db 421 TGTGACGATCAGCGATCTTACTTCTGCTATGACCACTGAGGCTTCTGTGTGG 480

QY 555 TGGAGCCCTGACCTCTGACCTCTACGACCCATGTGTGGCTGGCGCGCTGGCGCGCTGGG 614
Db 481 TGGAGCCCTGACCTCTGACCTCTACGACCCATGTGTGGCTGGCGCGCTGGCGCGCTGGG 540

QY 615 CACCGTGATCTGTATCAGCCCTTGGAGCTTATGGGACAAAGCTCAGGCTCAGCATGT 674
Db 541 CACCGTGATCTGTATCAGCCCTTGGAGCTTATGGGACAAAGCTCAGGCTCAGCATGT 600

QY 675 GTCTGACGGAGCTGGGTGCTGTGTTGCACTGACGTGCTCAGGCTGGCGGCTC 734
Db 601 GTCTGACGGAGCTGGGTGCTGTGTTGCACTGACGTGCTCAGGCTGGCGGCTC 660

QY 735 ACTGTGGCTGGGCTGGGCGCCCACTGCCCTTCGAGATGCGCTTCTCAGCGCTGTACTG 794
Db 661 ACTGTGGCTGGGCTGGGCGCCCACTGCCCTTCGAGATGCGCTTCTCAGCGCTGTACTG 720

QY 795 GTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGTTTCAAGGCGGAGGACGACATTC 854
Db 721 GTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGTTTCAAGGCGGAGGACGACATTC 780

QY 855 TGTGGGATGAGCTTTTGGCTGGTGGCATCTCAGGACCGTGGCTGCTGCTGCTCT 914
Db 781 TGTGGGATGAGCTTTTGGCTGGTGGCATCTCAGGACCGTGGCTGCTGCTGCTCT 840

QY 915 ACCCTTTGAGCTGTAAAGACCCCAAGCCAGGCTGGCTCTCGGA 957
Db 841 ACCCTTTGAGCTGTAAAGACCCCAAGCCAGGCTGGCTCTCGGA 884

RESULT 8
B0925174
LOCUS B0925174
DEFINITION AGENCOURT_8819286 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6423038 5', mRNA sequence.
ACCESSION B0925174
VERSION B0925174.1 GI:22340205
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@phs-research.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2600 row: p column: 15
High quality sequence stop: 616.

FEATURES
source
1..932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6423038"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 154 a 294 c 290 g 201 t 3 others
ORIGIN

Query Match 56.2%; Score 869; DB 14; Length 932;
Best Local Similarity 97.7%; Pred. No. 8.2e-203;
Matches 911; Conservative 0; Mismatches 18; Indels 3; Gaps 3;

QY 538 AAGGCTTCTGTGTGGTGGAGCCCTGACCTCTGACCTTACGACCCCATGGTGGCTGGC 597
Db 1 AAGGCTTCTGTGTGGTGGAGCCCTGACCTCTGACCTTACGACCCCATGGTGGCTGGC 60

QY 598 GGGCTGGCCCTGGGACCGCTGACTGTGATCAGCCCTGGAGCTTATGGGACAAAG 657
Db 61 GGGCTGGCCCTGGGACCGCTGACTGTGATCAGCCCTGGAGCTTATGGGACAAAG 120

QY 658 CTGCGAGCTCAGCATGTGTGCTACCGGAGCTGGGTGGCTGTGTTCGAACTGCAGTGGCT 717

Db	121	CTGCAGGCTCAGCATGTGTGTTACCGGAGCTGGGTGCTGTCTTCGAACATCGATGGCT	180
QY	718	CAGGCTGGCTGGCGCTCAGCTGTGGCTGGGCTGGGCGCCCACTGCCCTTCGAGATGTGCCC	777
Db	181	CAGGCTGGCTGGCGCTCAGCTGTGGCTGGGCTGGGCGCCCACTGCCCTTCGAGATGTGCCC	240
QY	778	TTCTCAGCCCTGTAATCTGGTTCAATATAGAGCTGGTGAAGAGCTTGGCTCAATGGGTTCAGG	837
Db	241	TTCTCAGCCCTGTACTGGTTCAATATAGAGCTGGTGAAGAGCTGGCTCAATGGGTTCAGG	300
QY	838	CCGAGGACCAAGACTCTCTGTGGCATGAGCTTTGTGGCTGGTGGCATCTCAGGACGGTG	897
Db	301	CCGAGGACCAAGACTCTCTGTGGCATGAGCTTTGTGGCTGGTGGCATCTCAGGACGGTG	360
QY	898	GCTCAGTGTCTGACTCTACCCCTTTGACGTGGTAAAGACCCAAAGCCAGGTGCTCTCGGGA	957
Db	361	GCTCAGTGTCTGACTCTACCCCTTTGACGTGGTAAAGACCCAAAGCCAGGTGCTCTCGGGA	420
QY	958	GGCATTGAAGGCTGTGAGAGTGAACCCCTGTGATGTGAGCTTCACCTTGGCTGTGCTGGG	1017
Db	421	GGCATTGAAGGCTGTGAGAGTGAACCCCTGTGATGTGAGCTTCACCTTGGCTGTGCTGGG	480
QY	1018	AGGATCCGGGCCGAGTCGGSCACCAAGGAGCTCTTTTGCAGGCTTCCTTCCCTCGGATCATC	1077
Db	481	AGGATCCGGGCCGAGTCGGSCACCAAGGAGCTCTTTTGCAGGCTTCCTTCCCTCGGATCATC	540
QY	1078	AAGGCTGCCCCCTCCTGTGCCATCATGATCAGACACTATGAGTTGCJAAAAGCTTCTTC	1137
Db	541	AAGGCTGCCCCCTCCTGTGCCATCATGATCAGACACTATGAGTTGCJAAAAGCTTCTTC	600
QY	1138	CAGAGGCTGAACAGAGACGGCTTCTGGCGCGCTGAAGGGGCAAGGAGGCAAGACCCC	1197
Db	601	CAGAGGCTGAACAGAGACGGCTTCTGGCGCGCTGAAGGGGCAAGGAGGCAAGACCCC	660
QY	1198	GTCTCT-CCACGATGGGAGAGGCGAGGAGAGACCCAGGCCAAGTGCTTTTCTCTCAG	1256
Db	661	GTCTCTCCACAGATGGGAGAGGCGAGGAGGAGACCCAGGCCAAGTGCTTTTCTCTCAG	720
QY	1257	CAGTGAAGGAGGGGCTGTGTTTCCCTTCCCTCCCGGCGACAAGCTCCAGGGCAGGGCTGT	1316
Db	721	CAGTGAAGGAGGGGCTGTGTTTCCCTTCCCTCCCGGCGACNAGCTCCAGGGCAGGGCTGT	780
QY	1317	CCCTTGGGCGGCCAGCACTTCTCAGACACACTTCTTCTGCTGCTCCAGTGTGT-GG	1375
Db	781	CCCTTGGGCGGNCAGCACTTCTCAGACACACTTCTTCTGCTGCTCAANTGTGGG	840
QY	1376	GGATCATCACTTACCCACCCGCCAAGTTCAAGACCAATCTTCCAGCTGCCCCCTTCGTG	1435
Db	841	GGATCATCACTTACCCACCCGCCAGGTCAGGACCAANTCTTCCAGCTGCCCCCTTCGTG	900
QY	1436	TTTCCCTGT-CTTTTGTGTAGCTGGGCATGTC	1466
Db	901	TTTTCCTCTGGTTTGTGTGTAACTGGGCATGGC	932

RESULT 9
AL516722
LOCUS
DEFINITION
AL516722 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS09A007YA10 5
prime, mRNA sequence.
ACCESSION
AL516722 AL316722.1 GI:12780215
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 925)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
Contact: Genoscope
COMMENT
Genoscope - Centre National de Sequencage

```

BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 925
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODA007YA10"
/gene_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH103"
/notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng Liang Life technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 153 a 294 c 279 g 192 t 7 others
ORIGIN

Query Match 56.2%; Score 868.4; DB 9; Length 925;
Best Local Similarity 97.6%; Pred. No. 1.1e-202;
Matches 905; Conservative 6; Mismatches 13; Indels 3; Gaps 3;

QY 33 AGGCGGACGAGCTAATGGCGCGCTGGCCCTGGCGCGCGCCGACGAGCACCCAGC 92
Db 1 AGCGCGGAGAGCTAATGGCGCGCTGGCCCTGGCGCGCGCGCGCGAGCACCCAGC 60

QY 93 CTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACACGAGACCCCTGGCGGCATCAGCCCCCT 152
Db 61 CTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACACGAGACCCCTGGCGGCATCAGCCCCCT 120

QY 153 CCAGCAATGGTGGCTCAGGCACGGGGCTGTGGTTACTCTCTCTTCATGACACCCCT 212
Db 121 CCAGCAATGGTGGCTCAGGCACGGGGCTGTGGTTACTCTCTCTTCATGACACCCCT 180

QY 213 GGACGTGTGAAGTTTGGCTTGCAGTCTCAGCGCGCCTCCATGCGCAGCGAGCTGATGC 272
Db 181 GGAAGTGTGAAGTTTGGCTTGCAGTCTCAGCGCGCCTCCATGCGCAGCGAGCTGATGC 240

QY 273 TTCTCCAGACTGTGAGGCTCTCCTATACAAATGGAAGTGCCTTCCTGTTATGCAATGG 332
Db 241 TTCTCCAGACTGTGAGGCTCTCCTATACAAATGGAAGTGCCTTCCTGTTATGCAATGG 300

QY 333 TGTCTCGAGCTCTGTAAGTGTGCGCAATGTGCGCGCTGTGCGCACTGGTTTCAGA 392
Db 301 TGTCTCGAGCTCTGTAAGTGTGCGCAATGTGCGCGCTGTGCGCACTGGTTTCAGA 360

QY 393 CCTTACCCGCTTCACTGGCACCATGATGCTTCGTGAAGATCGTGAGGCACGAGGGCAC 452
Db 361 CCTTACCCGCTTCACTGGCACCATGATGCTTCGTGAAGATCGTGAGGCACGAGGGCAC 420

QY 453 CAGGACCTCTGAGGCGGCTCCCGCGCAACCTGGTGATGATGTGCGCAGTACCGCAT 512
Db 421 CAGGACCTCTGAGGCGGCTCCCGCGCAACCTGGTGATGATGTGCGCAGTACCGCAT 480

QY 513 CTACTTCAGCTATGACCAACTGAAGGCTTTCGTGTGTGTGAGGCTTGACCTCTGCA 572
Db 481 CTACTTCAGCTATGACCAACTGAAGGCTTTCGTGTGTGTGAGGCTTGACCTCTGCA 540

QY 573 CCTCTACGACCATGGTGGCTGGCGCGCTGGCGCCCTGGGCACCGTGATGTGATCAG 632
Db 541 CCTCTACGACCATGGTGGCTGGCGCGCTGGCGCCCTGGGCACCGTGATGTGATCAG 600

QY 633 CCCCTGAGGCTATCGGCAAAAGCTGAGGCTCAAGATGTGTGCTGACGGAGCTGGG 692
Db 601 CCCCTGAGGCTATCGGCAAAAGCTGAGGCTCAAGATGTGTGCTGACGGAGCTGGG 660

```


QY 477 CGCCACCTCTGTGATGATCTGTGCCAGCTACCGGATCTACTTCACTGCTGCTATGACCAACT 536
 DB 484 CGCCACCTCTGTGATGATCTGTGCCAGCTACCGGATCTACTTCACTGCTGCTATGACCAACT 543
 QY 537 GAAGGCTTCTGTGCTGCGAGCTGACCTGACCTCTACGACCTTACGACCCATGCTGCTGG 596
 DB 544 GAAGGCTTCTGTGCTGCGAGCTGACCTGACCTCTACGACCTTACGACCCATGCTGCTGG 603
 QY 597 CGCGTGGCCCTGCGGACCTGACCTGATGATGACGCGCCCTGAGCTTATCGGACAAA 656
 DB 604 CGCGTGGCCCTGCGGACCTGACCTGATGATGACGCGCCCTGAGCTTATCGGACAAA 663
 QY 657 GCTGAGGCTCAGCATGTGCTACCGGAGCTGGGCTGCTGTTGCACTGCACTGCACTGGC 716
 DB 664 GCTGAGGCTCAGCATGTGCTACCGGAGCTGGGCTGCTGTTGCACTGCACTGCACTGGC 723
 QY 717 TCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 776
 DB 724 TCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 783
 QY 777 CTTCTACGCTCTGATGCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 836
 DB 784 CTTCTACGCTCTGATGCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 843
 QY 837 GCGGAGGACGACGACTTCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 894
 DB 844 GCGGAGGACGACGACTTCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 903
 QY 895 GTGGCTGAGCTGCTGACTTACCTTTGACGTGGTA 930
 DB 904 GTGGCTGAGCTGACTTACCTTTGACGGGGA 939

RESULT 14
 BQ934841
 LOCUS
 DEFINITION AGENCOURT_8824927 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6423317
 5', mRNA sequence.
 ACCESSION BQ934841
 VERSION BQ934841.1 GI:22350224
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 965)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF/Gazdar
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2601 row: 1 column: 06
 High quality sequence stop: 626.
 Location/Qualifiers
 1..965
 /organism="Homo sapiens"
 /db xref="taxon:9606"
 /clone="IMAGE:6423317"
 /clone_lib="NIH_MGC_18"
 /tissue_type="large cell carcinoma"
 /lab_host="DHIO (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dr priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,

BASE COUNT 161 a 308 c 295 g 197 t 4 others
 ORIGIN
 Query Match 55.2%; Score 953.8; DB 14; Length 965;
 Best Local Similarity 95.4%; Pred. No. 4.5e-199;
 Matches 921; Conservative 0; Mismatches 36; Indels 8; Gaps 4;
 QY 1 GCGTAGGTGCGCTCGAGCGCGG----GAGCCAGGCGGCGACGAGTAAATGGGCC 56
 DB 1 GCGTAGGTGCGCTCGAGCGCGGCGGACCGCGACAGCGCGCGGATATGGGCC 60
 QY 57 GCGTGGCCCTGGGCGCGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 116
 DB 61 GCGTGGCCCTGGGCGCGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACG 120
 QY 117 GATGCTGACGAGGACCTCGGGGATCAGCCCCCTCAGGAAATGCTGGCTCAGGAC 176
 DB 121 GATGCTGACGAGGACCTCGGGGATCAGCCCCCTCAGGAAATGCTGGCTCAGGAC 180
 QY 177 CGGGGCTGTGGTACCTCTCTCTCATGACACCCCTGACGCTGCTGAAGTTTCGCTGCA 236
 DB 181 CGGGGCTGTGGTACCTCTCTCTCATGACACCCCTGACGCTGCTGAAGTTTCGCTGCA 240
 QY 237 GTCTCAGCGGCCCTCCATGACGAGGCTGATGCTTCCTCAGAGCTGTGAGGCTCTC 296
 DB 241 GTCTCAGCGGCCCTCCATGACGAGGCTGATGCTTCCTCAGAGCTGTGAGGCTCTC 300
 QY 297 CTATACCAATGGAAGTGCCTCTCTGATTTGCAAGCTTCACTGACCTTCACTGGACCAT 356
 DB 301 CTATACCAATGGAAGTGCCTCTCTGATTTGCAAGCTTCACTGACCTTCACTGGACCAT 360
 QY 357 CCCAAATGCTGCGCGCTGTGCCACCTTGTTCAGAGCCCTACCCGCTTCACTGGACCAT 416
 DB 361 CCCAAATGCTGCGCGCTGTGCCACCTTGTTCAGAGCCCTACCCGCTTCACTGGACCAT 420
 QY 417 GATGCTTCTGTGAAGATCGTGAAGGACGAGGCGACGAGGACCTCTGAGCGGCTCTCC 476
 DB 421 GATGCTTCTGTGAAGATCGTGAAGGACGAGGCGACGAGGACCTCTGAGCGGCTCTCC 480
 QY 477 GCGCACCTGCTGATGACTGCGGACCTACCGCCATCTACTTCACTGCTGCTGACCACT 536
 DB 481 GCGCACCTGCTGATGACTGCGGACCTACCGCCATCTACTTCACTGCTGCTGACCACT 540
 QY 537 GAAGGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
 DB 541 GAAGGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 597 CGGCTGGCCCTCGGCGACCGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 656
 DB 601 CGGCTGGCCCTCGGCGACCGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 657 GCTGAGGCTCAGCATGTGCTGACCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
 DB 661 GCTGAGGCTCAGCATGTGCTGACCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 717 TCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 774
 DB 721 TCAGGCTGGCTGGCGCTCACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 780
 QY 775 CCCTTTCAGCCCTGTACTGGTTCACATATGA-GCTGGTGAAGAGCTGGCTCAATGGCTT 833
 DB 781 CCCTTTCAGCCCTGTACTGGNNTCACTATGAGGCTGGTGAAGAAGCTGCTCAATGGCTT 840
 QY 834 CAGCGCGAAGGACGACGACTTCTGTGGCATGAGC-TTGTGGCTGGTGGCATCTCAGGGA 892
 DB 841 CAGNCGAAGAACAGACTTCTGTGGCATGAGCTTTTGTGGCTGGGCGCATCTCAGGGA 900
 QY 893 CGGTGGCTCAGTGTGCTGACTCTACCCCTTTGAGTGGTAAAGACCCAAACGCCAGGTCCCTC 952
 DB 901 CGGTGGCTCAGTGTGCTGACTCTACCCCTTTGAGTGGTAAAGACCCAAACGCCAGGTCCCTC 960

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 19:04:17 ; Search time 3022.46 Seconds
(without alignments)
10726.543 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagcttcaagatggctg.....aggggcaaggaggaaggac 1114

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
To: number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :

- 1: gb_ba:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pi:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pi:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_ror:*
- 36: em_htg_ran:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1114	100.0	1114	6	AX403084 Sequence
2	1114	100.0	1114	9	AF317711 Homo sapi
3	1056	94.8	1550	9	HS801436 Homo sapi
4	1056	94.8	1560	9	BC009330 Homo sapi
5	1054.4	94.6	1546	9	AX403085 Sequence
6	1054.4	94.6	1556	9	AF151827 Homo sapi
7	1054.4	94.6	1579	9	BC001398 Homo sapi
8	1054.4	94.6	1643	6	AR116038 Sequence
9	1052.8	94.5	1554	9	AK026060 Homo sapi
10	1002.4	90.0	1402	17	AF119864 Homo sapi
11	651	58.4	173434	2	AC098934 Homo sapi
12	560.8	50.3	128802	2	AC123139 Rattus no
13	558.8	50.2	55252	2	AC119489_3 Continuation (4 of
14	554	49.7	156175	2	AC111356 Rattus no
15	552.8	49.6	195701	2	AC124203 Rattus no
16	488.8	43.9	177613	2	AC123014 Rattus no
17	472	42.4	704	6	AX305924 Sequence
18	448.4	40.3	758	6	AX421524 Sequence
19	440	39.5	194025	2	AC114721 Rattus no
20	298.4	26.8	138839	2	AC108295 Rattus no
21	242.2	21.7	1284	9	BC027322 Homo sapi
22	240.6	21.6	1517	9	AF125531 Homo sapi
23	210.4	18.9	162456	2	AC099097 Rattus no
24	176.4	15.8	139488	9	AC003043 Homo sapi
25	176.4	15.8	180520	2	AC019152 Homo sapi
26	175.2	15.7	195768	2	AC068775 Homo sapi
27	174.8	15.7	55774	2	AC091303 Homo sapi
28	170.2	15.3	512	6	AX400653 Sequence
29	170	15.3	537	6	AX106606 Sequence
30	170	15.3	537	6	AX148897 Sequence
31	170	15.3	537	6	AX200757 Sequence
32	170	15.3	537	6	AX267413 Sequence
33	168.2	15.1	82586	2	AC096040 Rattus no
34	168.2	15.1	188670	2	AC127606 Rattus no
35	156	14.0	194025	2	AC114721 Rattus no
36	142	12.7	189391	2	AC019151 Homo sapi
37	140.4	12.6	165490	2	AC025326 Homo sapi
38	140.4	12.6	190430	2	AC087873 Mus muscu
39	140.4	12.6	219037	10	AL596258 Mouse DNA
40	139.6	12.5	197197	2	AC118439 Rattus no
41	136.4	12.2	217422	2	AC116713 Mus muscu
42	131	11.8	1070	3	AY070516 Drosophil
43	130.4	11.7	1463	3	AY051544 Drosophil
44	130.4	11.7	16791	2	AC018172 Drosophil
45	130.4	11.7	175041	3	AC008092 Drosophil

ALIGNMENTS

RESULT 1
AX403084
LOCUS AX403084 1114 bp DNA linear PAT 07-JUN-2002
DEFINITION Sequence 1 from Patent WO0198355.
ACCESSION AX403084
VERSION AX403084.1 GI:21388033
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Lewin, D., Adams, S.H. and Yu, X.X.
TITLE Cgi-69 compositions and methods of use
JOURNAL Patent: WO 0198355-A 1 27-DEC-2001;

FEATURES		Genentech, Inc. (US) ; Curagen Corporation (US)	
source	Location/Qualifiers		
BASE COUNT	194 a 349 c 332 g 239 t		
ORIGIN			
Query Match	100.0%; Score 1114; DB 6; Length 1114;		
Best Local Similarity	100.0%; Pred. No. 5.4e-244;		
Matches 1114; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 CTGAAGCTTCAAGATGGCTGACAGAGCCCTCGGGCATACAGCCCTCCAGCAATGGT 60		
Db	1 CTGAAGCTTCAAGATGGCTGACAGAGCCCTCGGGCATACAGCCCTCCAGCAATGGT 60		
QY	61 GGCCTCAGCAGCGGGCTGTGTACCTCTCTTCATGACACCCCTGGACGTGGTAA 120		
Db	61 GGCCTCAGCAGCGGGCTGTGTACCTCTCTTCATGACACCCCTGGACGTGGTAA 120		
QY	121 GGTGGCTTGAAGTCTCAGCGCCCTCCATGCGCAGCAGTGTATGCTTCTTCAGACT 180		
Db	121 GGTGGCTTGAAGTCTCAGCGCCCTCCATGCGCAGCAGTGTATGCTTCTTCAGACT 180		
QY	181 GTGGAGCTCTCTATACCAAAATGGCCCTCTCTTCCAAATCCACAGGAAGTGCCTCT 240		
Db	181 GTGGAGCTCTCTATACCAAAATGGCCCTCTCTTCCAAATCCACAGGAAGTGCCTCT 240		
QY	241 GTATTCGAATGTGCTCGAGGCTCTGTACTGTGCCCAATGTGCCCGCTGTCCAC 300		
Db	241 GTATTCGAATGTGCTCGAGGCTCTGTACTGTGCCCAATGTGCCCGCTGTCCAC 300		
QY	301 CTGGTTTCAAGACCCCTACCCGCTTCACTGGCACCATTGGATGCTTCGTGAAGTGTGAG 360		
Db	301 CTGGTTTCAAGACCCCTACCCGCTTCACTGGCACCATTGGATGCTTCGTGAAGTGTGAG 360		
QY	361 GACAGAGGACACAGGACCCCTCGAGCGGCTCCCGCCACCTCGTGTGATGTGCC 420		
Db	361 GACAGAGGACACAGGACCCCTCGAGCGGCTCCCGCCACCTCGTGTGATGTGCC 420		
QY	421 AGCTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTTGTTGCTGAGC 480		
Db	421 AGCTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTTGTTGCTGAGC 480		
QY	481 CCTGACCTTGACCTTCTAGCACCCATGTGGCTGGCGGCTGGCCGCTGGGACCGT 540		
Db	481 CCTGACCTTGACCTTCTAGCACCCATGTGGCTGGCGGCTGGCCGCTGGGACCGT 540		
QY	541 GACTGTGATCGCCCTCGAGCTTATCGGACAAAGCTGACAGCTGACATGTGCTA 600		
Db	541 GACTGTGATCGCCCTCGAGCTTATCGGACAAAGCTGACAGCTGACATGTGCTA 600		
QY	601 CCGGAGCTGGGTGCTGTGTTGCAACTGACGAGTGTGCTGAGGCTGAGCTGCTG 660		
Db	601 CCGGAGCTGGGTGCTGTGTTGCAACTGACGAGTGTGCTGAGGCTGAGCTGCTG 660		
QY	661 GCTGGGTGGGGCCCACTGCCCCTTCAGATGTGCCCTTCTAGCCCTGTACTGTTCAA 720		
Db	661 GCTGGGTGGGGCCCACTGCCCCTTCAGATGTGCCCTTCTAGCCCTGTACTGTTCAA 720		
QY	721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCCAAGGACAGACTTCTGTGG 780		
Db	721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCCAAGGACAGACTTCTGTGG 780		
QY	781 CATGAGCTTTGTGGCTGGTGGATCTCAGGAGACGGTGGCTGAGTGTCACTACCTT 840		
Db	781 CATGAGCTTTGTGGCTGGTGGATCTCAGGAGACGGTGGCTGAGTGTCACTACCTT 840		
QY	841 TGAAGTGTGAAGACCAACCGCAGTGTCTCTGGAGCCATGAGGCTGTGAGTGA 900		
Db	841 TGAAGTGTGAAGACCAACCGCAGTGTCTCTGGAGCCATGAGGCTGTGAGTGA 900		
QY	901 CCCCTCGATGTGAGTCCACTGGCTGCTCGGAGGATCCGGCCAGTTCGGGCAC 960		
Db	901 CCCCTCGATGTGAGTCCACTGGCTGCTCGGAGGATCCGGCCAGTTCGGGCAC 960		

QY 61 GGCCTCAGGACCGGGCTGTGGTACCTCTCTCTCATGACACCCCTGGAGTGGTAA 120
 Db 61 GGCCTCAGGACCGGGCTGTGGTACCTCTCTCTCATGACACCCCTGGAGTGGTAA 120
 QY 121 GGTTCGCTCAGTCTCAGGGGCTCTCATGCGCAGGAGCTATGCTCTCTCAGACT 180
 Db 121 GGTTCGCTCAGTCTCAGGGGCTCTCATGCGCAGGAGCTATGCTCTCTCAGACT 180
 QY 181 GTGAGGCTCTCTATACAAATTCCTCTCTCTCAATCCACAGGGAATGCTCCT 240
 Db 181 GTGAGGCTCTCTATACAAATTCCTCTCTCTCAATCCACAGGGAATGCTCCT 240
 QY 241 GTATTGCAATGTGTCTCTGAGGCTCTGTACTGTGCCCCAATGCTGCGGCTGCGCAC 300
 Db 241 GTATTGCAATGTGTCTCTGAGGCTCTGTACTGTGCCCCAATGCTGCGGCTGCGCAC 300
 QY 301 CTGTTTCAGACCTACCGCTTCACTGSCACCATGGATGCTTCTGTAAGATCGTGG 360
 Db 301 CTGTTTCAGACCTACCGCTTCACTGSCACCATGGATGCTTCTGTAAGATCGTGG 360
 QY 361 GCACGAGGACAGGACCTCTGAGGCGGCTCTCCCGCCACCTGCTGATGATGCTGCC 420
 Db 361 GCACGAGGACAGGACCTCTGAGGCGGCTCTCCCGCCACCTGCTGATGATGCTGCC 420
 QY 421 AGCTACGCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGCTGCTGCGAGC 480
 Db 421 AGCTACGCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGCTGCTGCGAGC 480
 QY 481 CTTGACCTCTACCTGACGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 481 CTTGACCTCTACCTGACGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 541 GACTGTGATCAGCCCTCTGAGGCTTATCGGACCAAACTCGAGGCTCAGCATGTGCTGA 600
 Db 541 GACTGTGATCAGCCCTCTGAGGCTTATCGGACCAAACTCGAGGCTCAGCATGTGCTGA 600
 QY 601 CCGGAGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 601 CCGGAGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 661 CTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 CTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 CTATGAGCTGCTGAGAGCTGCTCAATGGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAG 780
 Db 721 CTATGAGCTGCTGAGAGCTGCTCAATGGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAG 780
 QY 781 CATGAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 CATGAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 TGAGCTGCTAAAGACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Db 841 TGAGCTGCTAAAGACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 901 CCCCCTCAGTGGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 901 CCCCCTCAGTGGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 961 CAAGGACTCTTTCAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 961 CAAGGACTCTTTCAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 1021 CATGATCAGCACCTATGAGTCTCGGCAAGCTTCTTCCAGAGGCTGAAACAGGACCGGCT 1080
 Db 1021 CATGATCAGCACCTATGAGTCTCGGCAAGCTTCTTCCAGAGGCTGAAACAGGACCGGCT 1080
 QY 1081 TCTGGGCTGAAAGGGCAAGGAGGCAAGGAC 1114
 Db 1081 TCTGGGCTGAAAGGGCAAGGAGGCAAGGAC 1114

RESULT 3
 HSM801436
 LOCUS
 DEFINITION Homo sapiens mRNA; cDNA DKF2p434C119 (from clone DKF2p434C119); complete cds.
 ACCESSION AL133584
 VERSION AL133584.1 GI:6599169
 KEYWORDS human
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1550)
 AUTHORS Wiemann,S., Weil,B., Wellenreuther,R., Gassenhuber,J., Glassl,S., Ansorho,W., Boecker,M., Bloeker,H., Bausersachs,S., Blum,H., Lauber,J., Duesterhoeft,A., Seyer,A., Koehrer,K., Strack,N., Meves,H.W., Ottenwaelder,B., Obermaier,B., Tampe,J., Heubner,D., Wambutt,R., Korn,B., Klein,M. and Poustka,A.
 TITLE Toward a Catalog of Human Genes and Proteins: Sequencing and Analysis of 500 Novel Complete Protein Coding Human cDNAs
 JOURNAL Genome Res. 11 (3), 422-435 (2001)
 PUBMED 11230166
 REFERENCE 2 (bases 1 to 1550)
 AUTHORS Poustka,A., Klein,M., Meves,H.W., Gassenhuber,J. and Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (15-DEC-1999) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKF2p434C119) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clonesrzd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.
 FEATURES
 Location/Qualifiers
 1..1550
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /clone="DKF2p434C119"
 /tissue_type="testis"
 /clone_lib="434 (synonym: htes3). Vector pSporti; host Dna10B; sites NotI + SalI"
 /dev_stage="adult"
 118..1173
 /gene="DKF2p434C119"
 118..1173
 /gene="DKF2p434C119"
 /note="CGI-69 protein, strong similarity to mitochondrial carrier proteins"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB63728.1"
 /db_xref="GI:6599170"
 /db_xref="SPTREMBL:Q9UF66"
 /translation="MADQDPAGISPLQQMVASGTGAVVTSLFMTPLDVKVRLQSQRP SVKIMFSSRLWSLYTKWCLLYCNGVLEPLCPNGARCATFQDPTFRFTGTMDA FYKIVRHGRTLMGSLPATLVMTVPATAIYFTAYDOLKAPLCGRALTSLYAPMVAG ALARGTIVTIVSPLLMRTKLOAHVSYRELGCACVRTAAGGWRSLWLGWPTALRD VPFSALYFNVELVKSLNGLRPEKQTSVGMGSFVAGISGTVAVLTLPDFVVKTRQK VALGMEAVRVNPLVDSTWLLLRIRAESGTGKLFAGFUPRTITKAAPSCAIMISTYE FKSFQRLNQDRLGG"
 polyA_site 1540
 BASE COUNT 278 a 490 c 459 g 323 t
 ORIGIN
 Query Match 94.8%; Score 1056; DB 9; Length 1550;
 Best Local Similarity 97.8%; Pred. No. 9.3e-231;
 Matches 1090; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 CTGAAGCTTCAAGATGCTGATCCAGGACCTCTGGGATCAGCCCTCCAGCAATGGT 60
 Db 105 CTGAAGCTTCAAGATGCTGATCCAGGACCTCTGGGATCAGCCCTCCAGCAATGGT 164
 QY 61 GGCCTCAGGACCGGGCTGTGGTACTCTCTCTCATCACACCCCTGGACGTTGTA 120
 Db 165 GGCCTCAGGACCGGGCTGTGGTACTCTCTCTCATCACACCCCTGGACGTTGTA 224
 QY 121 GGTTCGCTGAGTCTCAGGACCGCTTCATGGCCAGGAGCTGATCCCTCCAGACT 180
 Db 225 GGTTCGCTGAGTCTCAGGACCGCTTCATGGCCAGGAGCTGATCCCTCCAGACT 284
 QY 181 GTGAGGCTCTCTTATACCAATTCGCTCTCTCTCAATCCAGGAGGAGTCCCTCT 240
 Db 285 GTGAGGCTCTCTTATACCAATTCGCTCTCTCTCAATCCAGGAGGAGTCCCTCT 320
 QY 241 GTATGCAATGATGCTGAGGACCTCTGATCTGTCGCAAAATGTCGCCCTCTGCCAC 300
 Db 321 GTATGCAATGATGCTGAGGACCTCTGATCTGTCGCAAAATGTCGCCCTCTGCCAC 380
 QY 301 CTGTTTCAAGACCTTACCGCTTCACTGGCAGCATGGATGCTTCGNGAGATCGTGAG 360
 Db 381 CTGTTTCAAGACCTTACCGCTTCACTGGCAGCATGGATGCTTCGNGAGATCGTGAG 440
 QY 361 GCACAGGCGCACGAGACCTCTGAGCGGCTCTCCCGCACCCCTGGTGTGATCTGTGCC 420
 Db 441 GCACAGGCGCACGAGACCTCTGAGCGGCTCTCCCGCACCCCTGGTGTGATCTGTGCC 500
 QY 421 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTCTCTGTGTGTGAGC 480
 Db 501 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTCTCTGTGTGTGAGC 560
 QY 481 CTTGACCTCTGACCTTACGACCATGCTGCTGCTGGGCGCTGGCCGCTGGCAGCGT 540
 Db 561 CTTGACCTCTGACCTTACGACCATGCTGCTGCTGGGCGCTGGCCGCTGGCAGCGT 620
 QY 541 GACTGTGATCAGCGCCCTGAGCTTATCGGACAAAGCTTCAGGCTCAGCTGTGTGCTA 600
 Db 621 GACTGTGATCAGCGCCCTGAGCTTATCGGACAAAGCTTCAGGCTCAGCTGTGTGCTA 680
 QY 601 CCGGAGCTGGTGTCTGTTGCAATGCTGAGTGGCTCAGGCTGGTGGCTCAGCTGTG 660
 Db 681 CCGGAGCTGGTGTCTGTTGCAATGCTGAGTGGCTCAGGCTGGTGGCTCAGCTGTG 740
 QY 661 GCTGGCTGGGCGCCCACTGCTTGGAGATGCTGCTGCTGAGCTGCTGCTGCTCAA 720
 Db 741 GCTGGCTGGGCGCCCACTGCTTGGAGATGCTGCTGCTGAGCTGCTGCTGCTCAA 800
 QY 721 CTATGAGCTGGTGAAGAGCTGCTCAATGGCTCAGGCGCAAGGACCACTTCTGTGG 780
 Db 801 CTATGAGCTGGTGAAGAGCTGCTCAATGGCTCAGGCGCAAGGACCACTTCTGTGG 860
 QY 781 CATGAGCTTGTGGCTGTGGATCTCAGGACCGTGGCTGAGTGTGCTGACTTACCTT 840
 Db 861 CATGAGCTTGTGGCTGTGGATCTCAGGACCGTGGCTGAGTGTGCTGACTTACCTT 920
 QY 841 TCAGTGTGAAGACCAAGCGGCTGCTGCTGGGAGCGATCGAGCTGTGAGAGTGA 900
 Db 921 TCAGTGTGAAGACCAAGCGGCTGCTGCTGGGAGCGATCGAGCTGTGAGAGTGA 980
 QY 901 CCCCCTGCTGTGACTTCACTGCTGCTGCTGCGGAGGATCCGGGCGGAGTGGGAC 960
 Db 981 CCCCCTGCTGTGACTTCACTGCTGCTGCTGCGGAGGATCCGGGCGGAGTGGGAC 1040
 QY 961 CAAGGACTTCTTTCAGGCTTCTCTCTCGGATCATCAAGGCTGCCCCCTCTGTGCCAT 1020
 Db 1041 CAAGGACTTCTTTCAGGCTTCTCTCTCGGATCATCAAGGCTGCCCCCTCTGTGCCAT 1100
 QY 1021 CATGATCAGACCTTATGAGTTCGGAAGGCTTCTTCAGAGCTGAACAGGACCGGCT 1080
 Db 1101 CATGATCAGACCTTATGAGTTCGGAAGGCTTCTTCAGAGCTGAACAGGACCGGCT 1160
 QY 1081 TCTGGGCGCTGAAAGGGGCAAGGAGCAAGGAC 1114

Db 1161 TCTGGGCGCTGAAAGGGGCAAGGAGCAAGGAC 1194

RESULT 4
 BC009330
 LOCUS
 DEFINITION
 Homo sapiens, clone MGC:16684 IMAGE:4126268, mRNA, complete cds.
 ACCESSION
 BC009330
 VERSION
 BC009330.1 GI:14424605
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1560)
 Strausberg,R.
 Direct Submission
 Submitted (12-JUN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-ramail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Shevchenko,Y., Wetheby,K.D., Beckstrom-Sternberg,S.M.,
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
 Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Snyder,B., Stantropop,S., Thomas,P.J.,
 Tionson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
 Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 26 Row: g Column: 12.

FEATURES
 source
 1..1560
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:16684 IMAGE:4126268"
 /tissue_type="Brain, neuroblastoma"
 /clone_lib="NIH MGC 19"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 103..1158
 /codon_start=1
 /product="Unknown (protein for MGC:16684)"
 /protein_id="AAH09330.1"
 /db_xref="GI:14424605"
 /translation="MADDPAGISPLQOMVASGTGAVVTSLEMTPLDVVKVRLQSORP
 SVASEMSSRLWSLSTTKWKLKLYCNGVLEPLYLCPNGARCATFQDPTRTGTMDA
 FVALVRHGSTRILWSGFATLVMTVPATAYFTADQKAFGLGRALTSGLVAPVAG
 AKALGTVTIVSPLKMRKTLQHVSRVRELACVATAGQGWRSRLWGLWGTALRD
 VPFSLYVFNVELVKSLNGLRPKDTSVGSFVAGISGTVAALTLPLPDVVKTQRQ
 VALGAEAVRVNPLHVDSTWLLLRIRAEBSGTGKGLFAGFLPRIIKAAPSCAIIISTYE
 FGKSPFORLNQDRLLGG"
 BASE COUNT 296 a 488 c 453 g 323 t
 ORIGIN
 Query Match 94.8%; Score 1056; DB 9; Length 1560;
 Best Local Similarity 97.8%; Pred. No. 9.3e-231;
 Matches 1090; Conservative 0; Mismatches 0; Indels 24; Gaps 1;
 QY 1 CTGAAGCTTCAAGATGCTGATCCAGGACCTCTGGGATCAGCCCTCCAGCAATGGT 60

Db 90 CTGAAGCTTCAAGATGGCTGACACAGGACCCCTGCGGCGCATCAGCCCTCCAGCAATGGT 149
C 61 GGCCTTCAGGACCGGGCTGTGTACCTCTCTCATGACACCCCTCGAGTGGTGAA 120
Db 150 GGCCTTCAGGACCGGGCTGTGTACCTCTCTCTCATGACACCCCTCGAGTGGTGAA 209
QY 121 GGTTCGCTGAGTCTCAGGGCCCTCCATGGCCAGGAGCTGATGCCCTCCAGACT 180
Db 210 GGTTCGCTGAGTCTCAGGGCCCTCCATGGCCAGGAGCTGATGCCCTCCAGACT 269
QY 181 GTGAGGCTTCTTATACCAATTTGGCCCTCTCTCTCAATCCACAGGAAATGCTCCT 240
Db 270 GTGAGGCTTCTTATACCAATTTGGCCCTCTCTCTCAATCCACAGGAAATGCTCCT 305
QY 241 GTATTGCAATGGTCTCTGAGCCTCTGTACTGTGCCCCAAATGCGCCGTGGCCAC 300
Db 306 GTATTGCAATGGTCTCTGAGCCTCTGTACTGTGCCCCAAATGCGCCGTGGCCAC 365
QY 301 CTGGTTTCAAGACCTTACCGGCTTCACTGGCACCATTGATGCCCTTGTGAAGATCGTAG 360
Db 366 CTGGTTTCAAGACCTTACCGGCTTCACTGGCACCATTGATGCCCTTGTGAAGATCGTAG 425
QY 361 GCACGAGGGCACAGGACCTCTGAGCGGCTCCGCGCCACCTGGTGAATGATGTGCC 420
Db 426 GCACGAGGGCACAGGACCTCTGAGCGGCTCCGCGCCACCTGGTGAATGATGTGCC 485
QY 421 AGCTACCGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTGTGTCGAGC 480
Db 486 AGCTACCGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTGTGTCGAGC 545
QY 481 CTTGACCTCTGACCTCTACGACCACTTGGTGGCTGGCGCTGGCCGCTGGGACCCGT 540
Db 546 CTTGACCTCTGACCTCTACGACCACTTGGTGGCTGGCGCTGGCCGCTGGGACCCGT 605
QY 541 GACTGTGATCAGCCCTGAGGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGTGTA 600
Db 606 GACTGTGATCAGCCCTGAGGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGTGTA 665
QY 601 CCGGAGCTGGTGGCTGTGTTGCAACTGCAGTGGCTCAGGCTGGTGGCTCAGCTGTG 660
Db 666 CCGGAGCTGGTGGCTGTGTTGCAACTGCAGTGGCTCAGGCTGGTGGCTCAGCTGTG 725
QY 661 GCTGGGCTGGGGCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
Db 726 GCTGGGCTGGGGCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 785
QY 721 CTATGAGCTGGTGAAGAGCTGAGTCAATGGGCTCAGGCGGAGGACAGACTTCTGTGG 780
Db 786 CTATGAGCTGGTGAAGAGCTGAGTCAATGGGCTCAGGCGGAGGACAGACTTCTGTGG 845
QY 781 CATGAGCTTTGTGGCTGGTGGATCTCAGGACGGTGGCTGCAGTGGCTGACTCTACCCCT 840
Db 846 CATGAGCTTTGTGGCTGGTGGATCTCAGGACGGTGGCTGCAGTGGCTGACTCTACCCCT 905
QY 841 TGAAGCTGTAAGAACCAAGCCAGTGGCTCTGGGAGCGATGAGGCTGTGAGAGTCAA 900
Db 906 TGAAGCTGTAAGAACCAAGCCAGTGGCTCTGGGAGCGATGAGGCTGTGAGAGTCAA 965
QY 901 CCCCCTGATGTGAGCTCCACCTGGCTGCTGCTGGGAGGATCCGGGCGAGTCGGGCAC 960
Db 966 CCCCCTGATGTGAGCTCCACCTGGCTGCTGCTGGGAGGATCCGGGCGAGTCGGGCAC 1025
QY 961 CAAGGACTCTTTTTCAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTCGCAT 1020
Db 1026 CAAGGACTCTTTTTCAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTCGCAT 1085
QY 1021 CATGATCAGACCTATGAGTGGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1080
Db 1086 CATGATCAGACCTATGAGTGGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1145
QY 1081 TCTGGGCGGCTGAAGGGCAGGAGCAAGGAC 1114

Db 1146 TCTGGCGGCTGAAAAAGGGCAAGGCAAGGAC 1179
RESULT 5
AX403085
LOCUS AX403085 1546 bp DNA linear PAT 07 JUN 2002
DEFINITION Sequence 2 from Patent WO0198355.
ACCESSION AX403085
VERSION AX403085.1 GI:21388034
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Lewin D., Adams S.H. and Yu X.X.
TITLE Cgi-69 compositions and methods of use
JOURNAL Patent: WO 0198355-A 2 27-DEC-2001;
Genentech, Inc. (US); Curagen Corporation (US)
FEATURES
source
1..1546
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 272 a 489 c 459 g 326 t
ORIGIN
Query Match 94.6%; Score 1054.4; DB 6; Length 1546;
Best Local Similarity 97.8%; Pred. No. 2.2e-230;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 1 CTGAAGCTTCAAGATGGCTGACACAGGACCCCTGCGGCGCATCAGCCCTCCAGCAATGGT 60
Db 105 CTGAAGCTTCAAGATGGCTGACACAGGACCCCTGCGGCGCATCAGCCCTCCAGCAATGGT 164
QY 61 GGCCTTCAGGACCGGGCTGTGTACCTCTCTCATGACACCCCTCGAGTGGTGAA 120
Db 165 GGCCTTCAGGACCGGGCTGTGTACCTCTCTCATGACACCCCTCGAGTGGTGAA 224
QY 121 GGTTCGCTGAGTCTCAGCGGCCCTCCATGGCCAGGAGCTGCTTCCCTCAGACT 180
Db 225 GGTTCGCTGAGTCTCAGCGGCCCTCCATGGCCAGGAGCTGCTTCCCTCAGACT 284
QY 181 GTGAGGCTTCTCTATACCAATTTGCCCTCTCTCTCCAAATCCACAGGAGGCTCTCT 240
Db 285 GTGAGGCTTCTCTATACCAATTTGCCCTCTCTCTCCAAATCCACAGGAGGCTCTCT 320
QY 241 GTATTGCAATGGTGTCTGAGGCTCTGTACTGTGCCAAAATGGTGGCGCTGTGCCAC 300
Db 321 GTATTGCAATGGTGTCTGAGGCTCTGTACTGTGCCAAAATGGTGGCGCTGTGCCAC 380
QY 301 CTGGTTTCAAGACCTTACCCGCTTCACTGGACCATGGATGCTTCTGTAAGATCGTGAG 360
Db 381 CTGGTTTCAAGACCTTACCCGCTTCACTGGACCATGGATGCTTCTGTAAGATCGTGAG 440
QY 361 GCACGAGGGCACAGGACCCCTGTGGAGGGCTTCCCGCCACCCCTGGTGAATGACTGTGCC 420
Db 441 GCACGAGGGCACAGGACCCCTGTGGAGGGCTTCCCGCCACCCCTGGTGAATGACTGTGCC 500
QY 421 AGCTACCGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTGTGTCGAGC 480
Db 501 AGCTACCGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTGTGTCGAGC 560
QY 481 CTTGACCTCTGAGCTCTACGACCCATGGTGGCTGGCGCTGGCCGCTGGGACCCGT 540
Db 561 CTTGACCTCTGAGCTCTACGACCCATGGTGGCTGGCGCTGGCCGCTGGGACCCGT 620
QY 541 GACTGTGATCAGCCCTTCGAGCTTATGCGGACAAAGCTCAGGCTCAGCATGTGTGCTGA 600
Db 621 GACTGTGATCAGCCCTTCGAGCTTATGCGGACAAAGCTCAGGCTCAGCATGTGTGCTGA 680
QY 601 CCGGAGCTGGGCTGCTGTGTCGAAGTGCAGTGGCTCAGGCTGGCTCAGGCTCAGCTGTG 660
Db 681 CCGGAGCTGGGCTGCTGTGTCGAAGTGCAGTGGCTCAGGCTGGCTCAGCTCAGCTGTG 740

QY	661	GCTGGGCTGGGCCCCCCTCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAA	720
Db	741	GCTGGGCTGGGCCCCCCTCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAA	800
QY	721	CTATGAGCTGGTGAAGAGCTGGCTCAATGGCTCAGGCGGAGGACGAGACTTCTGTGG	780
Db	801	CTATGAGCTGGTGAAGAGCTGGCTCAATGGCTCAGGCGGAGGACGAGACTTCTGTGG	860
QY	781	CATGAGCTTTTGGCTGGCTGGCTCAGGAGCGGTGGCTGCATGTGCTGACTACCCCTT	840
Db	861	CATGAGCTTTTGGCTGGCTGGCTCAGGAGCGGTGGCTGCATGTGCTGACTACCCCTT	920
QY	841	TGACGTGTTAAAGACCCAAAGCCAGGTGGCTCTGGAGCGGATGAGGCTGTGAGTGA	900
Db	921	TGACGTGTTAAAGACCCAAAGCCAGGTGGCTCTGGAGCGGATGAGGCTGTGAGTGA	980
QY	901	CCCCCTGCATGTGGACTCCACCTGGCTGCTGCGGAGGATCCGGCCGAGTGGGCGAC	960
Db	981	CCCCCTGCATGTGGACTCCACCTGGCTGCTGCGGAGGATCCGGCCGAGTGGGCGAC	1040
QY	961	CAAGGACTCTTTTGGAGCTTCTCTCGGATCATCAAGGCTGCCCTCTCTGTGGCAT	1020
Db	1041	CAAGGACTCTTTTGGAGCTTCTCTCGGATCATCAAGGCTGCCCTCTCTGTGGCAT	1100
QY	1021	CATGATCAGACCTATGAGTTGGGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGCT	1080
Db	1101	CATGATCAGACCTATGAGTTGGGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGCT	1160
QY	1081	TCTGGCGGCTCAAGGGGCAAGGAGGAGGAC	1114
Db	1161	TCTGGCGGCTCAAGGGGCAAGGAGGAGGAC	1194
RESULT 6			
AF151827			
Homo sapiens CGI-69 prote:n mRNA, complete cds.			
AF151827			
AF151827.1 GI:4929606			
Homo sapiens.			
Homo sapiens.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
Lai,C.H., Chou,C.Y., Ch'ang,L.Y., Liu,C.S. and Lin,W.			
Identification of novel human genes evolutionarily conserved in			
Caenorhabditis elegans by comparative proteomics			
Genome Res. 10 (5), 703-713 (2000)			
20272150			
10810093			
PUBMED			
Lin,W.-C.			
Direct Submission			
Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia			
Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan			
Location/Qualifiers			
1..1556			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/chromosome="17"			
118..1173			
/codon_start=1			
/product="CGI-69 protein"			
/protein_id="A034064.1"			
/db_xref="GI:4929607"			
/translation="MADQDAGISPLQMVASGTGAVVTSLEWTPLDVVKVLOSQRP			
SWAGELPSSRLWSLSYTKWKCLLYCNGVSEPLKCPNGGARGATWFDPTAFTGTMDA			
FYKIVRHEGRTLLWSGLPATLVMTVPATAIYFTAYDQLKAFICGLRGLTSDLYAPWMA			
ALARLGTVTIVISPLELMRKLQAHVSRVGLGACVTAQVAGGWSRLWLGWPTALRD			
VPPSALYWFNVLKSLNGFRPKDQTSVGMFSVAGGISGTVAAVLTPFDVVKTRQ			
VALGMENAVRNPLHVDSTWLLLRIRAEOSTKGLFAGLPIRIIKAAPSCAIIISTVE			
source			
cds			

BASE COUNT			
ORIGIN			
FKSPFQRLNQDRILGG"			
282 a 489 c 459 g 326 t			
Query Match			
94.8%; Score 1054.4; DB 9; Length 1556;			
Best Local Similarity 97.8%; Pred. No. 2.2e-230;			
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;			
QY	1	CTAAGCTTCAAGATGCTGACAGGACCTCGGGGATCAGCCCCCTCCAGCAATGCT	60
Db	105	CTAAGCTTCAAGATGCTGACAGGACCTCGGGGATCAGCCCCCTCCAGCAATGCT	164
QY	61	GGCTCAGGACACCGGGGCTGTGTACTCTCTTCAATGACACCCCTGGACCTGTGAA	120
Db	165	GGCTCAGGACACCGGGGCTGTGTACTCTCTTCAATGACACCCCTGGACCTGTGAA	224
QY	121	GGTTCCCTTCAGCTCTCAGGGGCTCCATGGCCAGGAGTGATGCTTCTCTCCAGCT	180
Db	225	GGTTCCCTTCAGCTCTCAGGGGCTCCATGGCCAGGAGTGATGCTTCTCTCCAGCT	284
QY	181	GTGGAGCTTCTCTATACCAAAATTGGCTCTCTCTCAATCCACAGGAGTGGCTCT	240
Db	285	GTGGAGCTTCTCTATACCAAAATTGGCTCTCTCTCAATCCACAGGAGTGGCTCT	320
QY	241	GTATTGCAATGTGTCTGAGGCTCTGTACTGTGCCCAAAATGTCGCCCTGTGCCAC	300
Db	321	GTATTGCAATGTGTCTGAGGCTCTGTACTGTGCCCAAAATGTCGCCCTGTGCCAC	380
QY	301	CTGGTTTCAAGACCTTACCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCGT	360
Db	381	CTGGTTTCAAGACCTTACCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCGT	440
QY	361	GCACGAGGACACAGGACCTCTGGAGGCGCTTCCGCCACCTTGTGTGATGCTGSC	420
Db	441	GCACGAGGACACAGGACCTCTGGAGGCGCTTCCGCCACCTTGTGTGATGCTGSC	500
QY	421	AGCTACCGGCTTACTTCTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGAGC	480
Db	501	AGCTACCGGCTTACTTCTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGAGC	560
QY	481	CCTGACCTCTGACCTTACGACCCATGGTGGCTGGCGGCTGGCGGCTGGCGACCGT	540
Db	561	CCTGACCTCTGACCTTACGACCCATGGTGGCTGGCGGCTGGCGGCTGGCGACCGT	620
QY	541	GACTGTGATCAGCCCCCTGGAGCTTATCGGCAAGAGCTCAGAGCTCAGATGTGCTGA	600
Db	621	GACTGTGATCAGCCCCCTGGAGCTTATCGGCAAGAGCTCAGAGCTCAGATGTGCTGA	680
QY	601	CCGGAGCTGGGTGCTGTGTTCGAACTGACGTGGCTCAGGGTGGCTGGCGCTCAGTGTG	660
Db	681	CCGGAGCTGGGTGCTGTGTTCGAACTGACGTGGCTCAGGGTGGCTGGCGCTCAGTGTG	740
QY	661	GCTGGGCTGGGGCCCACTGCCCCCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA	720
Db	741	GCTGGGCTGGGGCCCACTGCCCCCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA	800
QY	721	CTATGAGCTGGTGAAGAGCTGCTCAATGGGCTCAGCGGAGGACCAAGCTTCTGTGGG	780
Db	801	CTATGAGCTGGTGAAGAGCTGCTCAATGGGCTCAGCGGAGGACCAAGCTTCTGTGGG	860
QY	781	CATGAGCTTTTGGCTGGTGGCATCTCAGGAGCGGTGGCTGCGAGTCTGACTTACCCCT	840
Db	861	CATGAGCTTTTGGCTGGTGGCATCTCAGGAGCGGTGGCTGCGAGTCTGACTTACCCCT	920
QY	841	TGACGTGTTAAAGACCCAAAGCCAGGTGGCTCTGGAGCGGATGAGGCTGTGAGTGA	900
Db	921	TGACGTGTTAAAGACCCAAAGCCAGGTGGCTCTGGAGCGGATGAGGCTGTGAGTGA	980
QY	901	CCCCCTGCATGTGGACTCCACCTGGCTGCTGCGGAGGATCCGGCCGAGTGGGCGAC	960
Db	981	CCCCCTGCATGTGGACTCCACCTGGCTGCTGCGGAGGATCCGGCCGAGTGGGCGAC	1040
QY	961	CAAGGACTCTTTTGGAGCTTCTCTCGGATCATCAAGGCTGCCCTCTCTGTGGCAT	1020

Db 1041 CAAGGAGCTCTTTCAGGCTTCTTCTCGATCATCAAGGCTGCCCCCTCTCTGTGCCAT 1100
 QY 1021 CATGATCAGCACCTATGAGTTGGGAAAAGCTTCTTCCAGAGGCTGAACACGAGCGGCT 1080
 Db 1101 CATGATCAGCACCTATGAGTTGGGAAAAGCTTCTTCCAGAGGCTGAACACGAGCGGCT 1160
 QY 1081 TCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
 Db 1161 TCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1194

RESULT 7
 BC001398 1579 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, CGI-69 protein, clone MGC:1163 IMAGE:3051436, mRNA,
 DEFINITION complete cds.
 ACCESSION BC001398
 VERSION BC001398.1 GI:12655092
 KEYWORDS MGC.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1579)
 Strausberg, R.
 Direct Submission
 Submitted (12-DEC-2000) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
 Tongson, E.E., Touchman, J.W., Tsurgeson, C., Vogt, J.L., Walker, W.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 4 Row: o Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10438782.
 Location/Qualifiers
 1. .1579
 /organism="Homo sapiens"
 /db_xref="locus:51629"
 /db_xref="taxon:9606"
 /clone="MGC:1163 IMAGE:3051436"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NIH_MGC_15"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 131. 1186
 /codon_start=1
 /product="CGI-69 protein"
 /protein_id="AAH01398.1"
 /db_xref="GI:12655093"
 /translation="MADQDPAGISPLQOMVASGTGAVVTSLFWTPLDVVKVLQSORP
 SMASLEMPSSRLWSLSYTKWKCLLYCNGVLEPLCPNGARCATWFDPRFTGTGND
 FVKIVRHGTRILWSGLPATLVMTVPATAIYFTAYDQLKAPLGRALTSLSLAPWVAG

FEATURES
 source

CDS

ALALGLGTVTISPFLMRETKLQAQHSVYRELGAQVETAVAGQWRSMLWGMPTALRD
 VPFSALYWNFYSLVMSWLNFRPKDQSVGVFVAGGIGSTVAVALTTPFPDWWKTQRO
 VALGAEAVRVNPLHVDSTWLLLRIRAEISGTLKGLFAGFLPRIIKAAFSCAIMISTYE
 FGKSFQRLNQDRLLGG"
 BASE COUNT 293 a 496 c 464 g 326 t
 ORIGIN

Query Match 94.6%; Score 1054.4; DB 9; Length 1579;
 Best Local Similarity 97.8%; Pred. No. 2.2e-230;
 Matches 1099; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACAGAACCTCGGGGCTATCAGCCCCCTCCAGCAAAATGCT 60
 Db 118 CTGAAGCTTCAAGATGGCTGACAGAACCTCGGGGCTATCAGCCCCCTCCAGCAAAATGCT 177
 QY 61 GGCCTCAGGACCGGGGCTGTGTATCTCTCTCTCATGACACCCCTGACGTGTGAA 120
 Db 178 GGCCTCAGGACCGGGGCTGTGTATCTCTCTCTCATGACACCCCTGACGTGTGAA 237
 QY 121 GGTTCGCTGCAGTCTCAGCGGCCCTCCATGGCCAGCGAGCTGATGCTTCTCCAGACT 180
 Db 238 GGTTCGCTGCAGTCTCAGCGGCCCTCCATGGCCAGCGAGCTGATGCTTCTCCAGACT 297
 QY 181 GTGGAGCTTCTCTATACCAAAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTGCCTCT 240
 Db 298 GTGGAGCTTCTCTATACCAAAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTGCCTCT 333
 QY 241 GTATTGCAATGCTCTCTGAGGCTCTGTACCTGTGCCAAATGGTCCCGCTGTGCCAC 300
 Db 334 GTATTGCAATGCTCTCTGAGGCTCTGTACCTGTGCCAAATGGTCCCGCTGTGCCAC 393
 QY 301 CTGGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCTGTAG 360
 Db 394 CTGGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCTGTAG 453
 QY 361 GCACGAC :CACGAGAACCTCTGAGCGGCTTCCCGCCACCTGTGTGATGACTGTGCC 420
 Db 454 GCACGAGGACACGAGACCTCTGAGCGGCTTCCCGCCACCTGTGTGATGACTGTGCC 513
 QY 421 AGCTACCGCATCTACTTCACTGCTATGACCACTGAAAGGCTTCTGTGTGTGCTGAGC 480
 Db 514 AGCTACCGCATCTACTTCACTGCTATGACCACTGAAAGGCTTCTGTGTGTGCTGAGC 573
 QY 481 CTTGACCTCTGACCTTACACACCATGTGTGCTGGCGCTGCGCCGCTGGGCGCTCACTGTG 540
 Db 574 CTTGACCTCTGACCTTACACACCATGTGTGCTGGCGCTGCGCCGCTGGGCGCTCACTGTG 633
 QY 541 GACTGTGATGAGCCCTCTGAGGCTTATGCGGACAAAGCTGACGCTCAGCATGTGTGCTGA 600
 Db 634 GACTGTGATGAGCCCTCTGAGGCTTATGCGGACAAAGCTGACGCTCAGCATGTGTGCTGA 693
 QY 601 CCGGGAGCTGGGCTGCTGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 694 CCGGGAGCTGGGCTGCTGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753
 QY 661 GCTGGGCTGGGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 754 GCTGGGCTGGGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
 QY 721 CTATGAGCTGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGACTTCTGTGGG 780
 Db 814 CTATGAGCTGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGACTTCTGTGGG 873
 QY 781 CATGAGCTTCTGTGCTGGGCTCTCAGGAGCGGTGGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 874 CATGAGCTTCTGTGCTGGGCTCTCAGGAGCGGTGGCTGCTGCTGCTGCTGCTGCTGCT 933
 QY 841 TCAGCTGTGAAGACCCCAACCGCAGCTGCTCTGGAGCGATGGAGGCTGTGAGAGTGAA 900
 Db 934 TCAGCTGTGAAGACCCCAACCGCAGCTGCTCTGGAGCGATGGAGGCTGTGAGAGTGAA 993
 QY 901 CCCCCTGATGTGAGCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960


```

Db 994 CCCCTCATGTGACTCCACCTGGCTGCTGCTGCTGGAGGATCGGGCGAGTCGGGCAC 1053
Qy 961 CAAGGACTCTTTGAGCTTCCCTCGATCATCAAGGCTGCCCCCTCTGTGTCAT 1020
Db 1054 CAAGGACTCTTTGAGCTTCCCTCGATCATCAAGGCTGCCCCCTCTGTGTCAT 1113
Qy 1021 CATGATCAGCACTATGATTCGCAAAAGCTTCTTCAGAGGCTGAACCAAGACGGCT 1080
Db 1114 CATGATCAGCACTATGATTCGCAAAAGCTTCTTCAGAGGCTGAACCAAGACGGCT 1173
Qy 1081 TCTGGCGGCTCAAGAGGGCAAGGAGGCAAGGAC 1114
Db 1174 TCTGGCGGCTCAAGAGGGCAAGGAGGCAAGGAC 1207

RESULT 8
AR116038 AR116038 1643 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 68 from patent US 6132973.
DEFINITION AR116038
ACCESSION AR116038
VERSION AR116038.1 GI:14096360
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Unclassified.
LAT, P., Hillman, J.L., Bandman, O., Shah, P., Au-Young, J., Yue, H.,
Guegler, K.J. and Corley, N.C.
TITLE Human regulatory molecules
JOURNAL Patent: US 6132973-A 68 17-OCT-2000;
FEATURES
Location/Qualifiers
source i..1643
/organism="unknown"
BASE COUNT 307 a 501 c 476 g 359 t
ORIGIN

Query Match 94.6%; Score 1054.4; DB 6; Length 1643;
Best Local Similarity 97.88; Pred. No. 2.1e-230;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

Qy 1 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCATCAGCCCTCCAGCAATGCT 60
Db 192 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCATCAGCCCTCCAGCAATGCT 251
Qy 61 GGCTCAGGACCGGGCTGTGTTACTCTCTTCATGACACCCCTGAGCATGTGAA 120
Db 252 GGCTCAGGACCGGGCTGTGTTACTCTCTTCATGACACCCCTGAGCATGTGAA 311
Qy 121 GGTTCGCTGAGTCTCAGCGGCTCCATGGCCAGGAGTGTCTTCTCCAGACT 180
Db 312 GGTTCGCTGAGTCTCAGCGGCTCCATGGCCAGGAGTGTCTTCTCCAGACT 371
Qy 181 GTGAGCTCTCTCTATACCAAAATGGCTCTCTTCCAATCCAGGAGTGTCTCT 240
Db 372 GTGAGCTCTCTCTATACCAAAATGGCTCTCTTCCAATCCAGGAGTGTCTCT 407
Qy 241 GTATTGCAATGTGTCTGAGGCTCTGTACTGTGCGCAATGTCGCCGCTGTGCCAC 300
Db 408 GTATTGCAATGTGTCTGAGGCTCTGTACTGTGCGCAATGTCGCCGCTGTGCCAC 467
Qy 301 CTGGTTCAAGACCTACCGCTTCACTGGCACCATGATGCTCTCTGAGATCTGTGAG 360
Db 468 CTGGTTCAAGACCTACCGCTTCACTGGCACCATGATGCTCTCTGAGATCTGTGAG 527
Qy 361 GCACGAGGACACGAGGACCTCTTGAGGCGCTCCCGCCACCCCTGGTGATGCTGTGCC 420
Db 528 GCACGAGGACACGAGGACCTCTTGAGGCGCTCCCGCCACCCCTGGTGATGCTGTGCC 587
Qy 421 AGCTACCGGCACTACTTCACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTCGAGC 480
Db 588 AGCTACCGGCACTACTTCACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTCGAGC 647
Qy 481 CCTGACCTCTGACCTTACGACACCATGCTGGTGGCTGGCGGCTGGCCGCTGGCCACT 540

```

```

Db 648 CTGACCTCTGACCTTACGACCCCATGTGCTGGCGCTGGCGCTGGCGCTGGCGACCGT 707
Qy 541 GACTGTGATCAGCCCTCTGGAGCTTATGCGGACAAAGCTGCGAGCTCAGCATGTGCGTA 600
Db 708 GACTGTGATCAGCCCTCTGGAGCTTATGCGGACAAAGCTGCGAGCTCAGCATGTGCGTA 767
Qy 601 CCGGG..CTGGGTGGCTGTGTTTGAACCTCAGTGGCTCAGGCTGGCTGGCGCTCACTGTG 660
Db 768 CCGGGAGCTGGGTGGCTGTGTTTGAACCTCAGTGGCTCAGGCTGGCTGGCGCTCACTGTG 827
Qy 661 GCTGGCTGGGGCCCCCCTGCTGCTTCGAGATGCTGCTTCACGCGCTGACTGTTCAA 720
Db 828 GCTGGCTGGGGCCCCCCTGCTGCTTCGAGATGCTGCTTCACGCGCTGACTGTTCAA 887
Qy 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGGG 780
Db 888 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGGG 947
Qy 781 CATGAGCTTTTGGCTGGTGGCTATCTCAGGACCGTGGCTCAGTGTGCTGCTTACCTT 840
Db 948 CATGAGCTTTTGGCTGGTGGCTATCTCAGGACCGTGGCTCAGTGTGCTGCTTACCTT 1007
Qy 841 TGACGTGTGAAGACCCCAACCGCAGGCTCTCTGCGAGCGATGGAGCTGTGAGATGAA 900
Db 1008 TGACGTGTGAAGACCCCAACCGCAGGCTCTCTGCGAGCGATGGAGCTGTGAGATGAA 1067
Qy 901 CCCCCTGCTATGAGCTCCACCTGGCTGCTGCTGGGAGGATCCGGCCCGAGTCGGGCAC 960
Db 1068 CCCCCTGCTATGAGCTCCACCTGGCTGCTGCTGGGAGGATCCGGCCCGAGTCGGGCAC 1127
Qy 961 CAAGGACTCTTTGAGGCTTCTCTTCTCGATCATCAAGGCTGCCCTCTCTGTGCTCAT 1020
Db 1128 CAAGGACTCTTTGAGGCTTCTCTTCTCGATCATCAAGGCTGCCCTCTCTGTGCTCAT 1187
Qy 1021 CATGATCAGCACTATGAGTTCGGCAAAAGCTTCTTTCAGAGGCTGAACCAAGACCGCT 1080
Db 1188 CATGATCAGCACTATGAGTTCGGCAAAAGCTTCTTTCAGAGGCTGAACCAAGACCGCT 1247
Qy 1081 TCTGGCGGCTCAAGAGGGCAAGGAGGCAAGGAC 1114
Db 1248 TCTGGCGGCTCAAGAGGGCAAGGAGGCAAGGAC 1281

RESULT 9
AK026050 AK026050 1554 bp mRNA linear PRI 29-SEP-2000
LOCUS Homo sapiens cDNA: FLJ22407 fis, clone HRC08407.
DEFINITION
ACCESSION AK026050
VERSION AK026050.1 GI:10438782
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens primary human renal epithelial cells cDNA to mRNA,
clone lib:HRC clone:HRC08407.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (sites)
Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Ohtani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1554)
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan [E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416]
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert

```



```

FH
FT source
FT 1. .1402
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="FLB8041"
FT /tissue_type="liver"
FT /dev_stage="fetus"
FT 4. .1017
FT /codon_start=1
FT /db_xref="SWISS-PROT:Q9BJJ4"
FT /note="Predicted protein of H2163"
FT /product="PR02163"
FT /protein_id="AAF9618.1"
FT /translation="WVAGSTGAVTSLFMTPLDVVKVLRQSRSMSELPSSRLWSL
FT SYTKWCKLLYKNGVLEPLYLCPNGARCAITWFDPTFRFTGTMDFVKIVRHEGRTILWSG
FT LPATLVMTVPATAYFTAYDOLKAFLCGRALTSDLYAPWAVGALARGTGTVTVISPLELM
FT RTLOACHVSYRELCAVCRTAVAOGGWRSLWLGWPTALRDVPRFSAVLYWENLYKSL
FT NGLRPKDOTSVGKMSFVAGGISRTVAALVTLTPDYVKTORVALGAMAEAVNPVHVDST
FT WLLRKRIRAESGTGLFAGFLPRIIKAPPSCAIMISTYEFKSFQRLNDRLLGG"
FT
XX
SQ Sequence 1402 bp; 261 A; 432 C; 401 G; 308 T; 0 other;

Query Match 90.0%; Score 1002.4; DB 17; Length 1402;
Best Local Similarity 97.6%; Pred. No. 1.6e-218;
Matches 1037; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 53 CAAATGTTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTTCATGACACCCCTGGAC 112
DB 1 CAAATGTTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTTCATGACACCCCTGGAC 60

QY 113 GTGGTGAAGGTTGCGCTGAGTCTCAGCGGCTCTCATGGCCAGGAGTGATGCTTCC 172
DB 61 GTGGTGAAGGTTGCGCTGAGTCTCAGCGGCTCTCATGGCCAGGAGTGATGCTTCC 120

QY 173 TCCAGACTGTGAGCTCTCTATACCAAAATGCCCTCTCTCTCCAAATCCACAGGGAAG 232
DB 121 TCAGACTGTGAGCTCTCTATACCAAAATGCCCTCTCTCTCCAAATCCACAGG 156

QY 233 TGCTCTCTGTATTTGAATGTGCTTGGAGCTCTGTACCTGTGCCCAAAATGGTGGCG 292
DB 157 TGCTCTCTGTATTTGAATGTGCTTGGAGCTCTGTACCTGTGCCCAAAATGGTGGCG 216

QY 293 TGTGCCACCTGTGTTCAAGACCTACCGCTTCACTGGCAGCTGATGCTTGTGAG 352
DB 217 TGTGCCACCTGTGTTCAAGACCTACCGCTTCACTGGCAGCTGATGCTTGTGAG 276

QY 353 ATCGTGAGGACGAGGACACAGACCCCTCTGAGCGGCTCTCCCGCCACCTGGTGATG 412
DB 277 ATCGTGAGGACGAGGACACAGACCCCTCTGAGCGGCTCTCCCGCCACCTGGTGATG 336

QY 413 ACTGTGCCAGCTACCGCATCTACTTCACTGCTGTATGACCAACTGAAGGCTTCTGTGT 472
DB 337 ACTGTGCCAGCTACCGCATCTACTTCACTGCTGTATGACCAACTGAAGGCTTCTGTGT 396

QY 473 GGTGAGGCTTGAACCTGTGACCTTACGACCATGCTGGCTGGCGGCTGGCGCCGCTG 532
DB 397 GGTGAGGCTTGAACCTGTGACCTTACGACCATGCTGGCTGGCGGCTGGCGCCGCTG 456

QY 533 GGCACCGTGACTGTGATCAGCGCCCTTATGCGGACAAAGCTGCGAGGCTCAGCAT 592
DB 457 GGCACCGTGACTGTGATCAGCGCCCTTATGCGGACAAAGCTGCGAGGCTCAGCAT 516

QY 593 GTGTGCTACCGGAGCTGGGTGCTGTGTTTGAATGTCAGTGGCTGAGGCTGCTGGCGC 652
DB 517 GTGTGCTACCGGAGCTGGGTGCTGTGTTTGAATGTCAGTGGCTGAGGCTGCTGGCGC 576

QY 653 TCAGTGTGCTGGGTGGGCGCCACTGCGCTTCAGAGTGTGCTTCTCAGCCCTGTAC 712
DB 577 TCAGTGTGCTGGGTGGGCGCCACTGCGCTTCAGAGTGTGCTTCTCAGCCCTGTAC 636

QY 713 TGGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGCAAGGACAGACT 772
DB 637 TGGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGCAAGGACAGACT 696

```

```

QY 773 TCTGTGGCATGAGCTTTGTGGCTGTGGCAITCTCAGGACGTTGGCTGCAGTGTGACT 832
DB 697 TCTGTGGCATGAGCTTTGTGGCTGTGGCAITCTCAGGACGTTGGCTGCAGTGTGACT 756
QY 833 CTACCTTTTACGTTGTAAGACCCAAACCCAGGTCGCTCTGGAGCGATGAGGCTGTG 892
DB 757 CTACCTTTTACGTTGTAAGACCCAAACCCAGGTCGCTCTGGAGCGATGAGGCTGTG 816
QY 893 AGAGTGAACCCCTGCACTGGACTCCACCTGCTGCTGCTGCGGAGATCCGGGCCGAG 952
DB 817 AGAGTGAACCCCTGCACTGGACTCCACCTGCTGCTGCTGCGGAGATCCGGGCCGAG 876
QY 953 TCGGACCAACAGGACTCTTTTGAGGCTTCTCTTCTCGGATCATCAAGGCTGCCCTTCC 1012
DB 877 TCGGACCAACAGGACTCTTTTGAGGCTTCTCTTCTCGGATCATCAAGGCTGCCCTTCC 936
QY 1013 TGTGCCATCATGATCAGACACTATGAGTTTCGCAAAAGCTTCTTCCAGAGGCTGAACAG 1072
DB 937 TGTGCCATCATGATCAGACACTATGAGTTTCGCAAAAGCTTCTTCCAGAGGCTGAACAG 996
QY 1073 GACCGCTTCTGGCGGCTGAAAGGGCCAGGAGCAAGGAC 1114
DB 997 GACCGCTTCTGGCGGCTGAAAGGGCCAGGAGCAAGGAC 1038

```

AC098934 173434 bp DNA linear HTG 06-NOV-2001
 Homo sapiens chromosome 1 clone RP11-480112, WORKING DRAFT
 SEQUENCE, 6 unordered pieces.
 AC098934 AL390563
 AC098934.1 GI:16751904
 HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
 Direct Submission
 Title
 JOURNAL
 Reference
 Authors
 Title
 JOURNAL
 Reference
 Authors
 Title
 JOURNAL
 Comment
 Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgctgs@u.washington.edu
 Drafting Center: SC
 ----- Project Information -----
 Center project name: chr-1
 Center clone name: RP11-480112 (sc0719)
 ----- Summary Statistics -----
 Sequencing vector: plasmid; 32% of reads
 Sequencing method: 108752; 68% of reads
 Chemistry: Dye-terminator ET; 83% of reads
 Chemistry: Dye-terminator Big Dye; 11% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 170284 bases at least Q40
 Consensus quality: 171568 bases at least Q30
 Consensus quality: 172298 bases at least Q20
 Insert size: 172934; sum-of-contigs
 Quality coverage: 8.7x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

Correll, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsone, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, J. C., Lewis, L.,
Li, J., Li, Z., Lichtenberg, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, R., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawiney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogah, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

4632 5802: contig of 1171 bp in length
5803 5902: gap of unknown length
5903 7060: contig of 1158 bp in length
7061 7160: gap of unknown length
7161 8514: contig of 1354 bp in length
8515 8614: gap of unknown length
8616 9672: contig of 1058 bp in length
9673 9772: gap of unknown length
9773 12180: contig of 2408 bp in length
12181 12280: gap of unknown length
12281 13463: contig of 1183 bp in length
13464 13563: gap of unknown length
13564 16106: contig of 2543 bp in length
16107 16206: gap of unknown length
16207 18884: contig of 2678 bp in length
18885 18984: gap of unknown length
18986 20532: contig of 1548 bp in length
20533 20632: gap of unknown length
20633 23210: contig of 2578 bp in length
23211 23310: gap of unknown length
23311 25515: contig of 2205 bp in length
25516 25615: gap of unknown length
25616 27411: contig of 1796 bp in length
27412 27511: gap of unknown length
27512 29449: contig of 1938 bp in length
29450 32307: gap of unknown length
32308 32407: gap of unknown length
32408 35335: gap of unknown length
35336 36153: contig of 3518 bp in length
36154 39253: gap of unknown length
39254 43671: contig of 4418 bp in length
43672 43771: gap of unknown length
43772 47082: contig of 3311 bp in length
47083 47182: gap of unknown length
47183 52198: contig of 5016 bp in length
52199 52981: gap of unknown length
52982 57787: contig of 3989 bp in length
57788 57887: gap of unknown length
57888 60500: contig of 4113 bp in length
60501 60600: gap of unknown length
60601 66994: contig of 6394 bp in length
66995 67094: gap of unknown length
67095 73183: contig of 6089 bp in length
73184 73283: gap of unknown length
73284 80968: contig of 7685 bp in length
80969 81068: gap of unknown length
81069 98060: contig of 16992 bp in length
98061 98160: gap of unknown length
98161 128802: contig of 30642 bp in length.

FEATURES

source

1..128802
/organism="Rattus norvegicus"
/db_xref="taxon:101116"
/clone="CH230-11K3"

BASE COUNT 21721 a 38987 c 24403 g 38221 t 5470 others

ORIGIN

Query Match 50.3%; Score 560.8; DB 2; Length 128802;
Best Local Similarity 79.0%; Fred. No. 7.9e-118;
Matches 719; Conservative 0; Mismatches 182; Indels 9; Gaps 4;

Qy 2 TGAAGCTTCAAGATGGCTGACACGAGACCTCGGGGCATCAGCCCTCCAGCAATGGTG 61
Db 78069 TGAATTCACAGATGGTTGACAGAGATCTCTGGTGCATTAGACCTTTTACACATGGTG 78010
Qy 62 GCCTCAGGCACCGGGGCTGTGGTTACCTCTCTCTCATGACACCCCTCGACCTGTGAG 121
Db 78009 GCCTCAGGAACCTGGGGTGTGGTGCATCTCCCTTTCATGACACCCCTGTGTAGTGAAG 77950
Qy 122 GTTCGCTCGAGTCTCA-----CGGCCCTCCATGGCCGAGAGCTGATGCTCTCTCCA 176

RESULT 14

AC111356

LOCUS

DEFINITION

AC111356

VERSION

KEYWORDS

SOURCE

ORGANISM

AC111356 156175 bp DNA linear HTG 13-JUL-2002
Rattus norvegicus clone CH230-161G12, *** SEQUENCING IN PROGRESS
***, 66 unordered pieces.

AC111356.2 GI:21735881

HTG; HTGS PHASE1.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 156175)

REFERENCE

AUTHORS

Murny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayela,M., Banks,T.,
Barbacia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.B., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Dunn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,K., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elnaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J.J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtenage,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nckenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojboomani,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooftari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Tang,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Thomas,N., Thomas,S.,
Tansey,J., Taylor,A., Taylor,T., Telford,B., Thomas,N., Wang,Q.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleciyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission

2 (bases 1 to 156175)

Worley,K.C.

Direct Submission

Submitted (19-FEB-2002)

Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 156175)

Worley,K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Jul 12, 2002 this sequence version replaced gi:18701120.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GMCE

Center clone name: CH230-161G12

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990325

Consensus quality: 99973 bases at least Q40

Consensus quality: 106615 bases at least Q30

Consensus quality: 111100 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 66 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will

* be preserved.

* 1 1025: contig of 1025 bp in length

* 1026 1125: gap of unknown length

* 1126 2218: contig of 1093 bp in length

* 2219 2318: gap of unknown length

* 2319 3454: contig of 1136 bp in length

* 3455 3554: gap of unknown length

* 3555 4690: contig of 1136 bp in length

* 4691 4790: gap of unknown length

* 4791 5955: contig of 1165 bp in length

* 5956 6085: gap of unknown length

* 6086 7439: contig of 1384 bp in length

* 7440 7539: gap of unknown length

* 7540 8707: contig of 1168 bp in length

* 8708 8807: gap of unknown length

* 8808 10033: contig of 1226 bp in length

* 10034 10133: gap of unknown length

* 10134 11778: contig of 1645 bp in length

* 11779 11878: gap of unknown length

* 11879 13633: contig of 1755 bp in length

* 13634 13733: gap of unknown length

* 13734 14947: contig of 1214 bp in length

* 14948 15047: gap of unknown length

* 15048 15659: contig of 1522 bp in length

* 15660 16669: gap of unknown length

* 16670 17786: contig of 1117 bp in length

* 17787 17886: gap of unknown length

* 17887 19738: contig of 1852 bp in length

* 19739 19838: gap of unknown length

* 19839 21532: contig of 1694 bp in length

* 21533 22632: gap of unknown length

* 22633 22715: contig of 1083 bp in length

* 22716 22815: gap of unknown length

* 22816 24052: contig of 1237 bp in length

* 24053 24152: gap of unknown length

* 24153 25250: contig of 1098 bp in length

* 25251 25944: contig of 1594 bp in length

* 25945 27044: gap of unknown length

* 27045 28474: contig of 1430 bp in length

* 28475 28574: gap of unknown length

* 28575 29825: contig of 1351 bp in length

* 29826 30025: gap of unknown length

* 30026 31402: contig of 1377 bp in length

* 31403 31502: gap of unknown length

* 31503 33645: contig of 2143 bp in length

* 33646 33745: gap of unknown length

* 33746 35290: contig of 1545 bp in length

* 35291 35390: gap of unknown length

* 35391 36672: contig of 1282 bp in length

* 36673 36772: gap of unknown length

* 36773	38565:	contig of 1793 bp in length	
* 38566	38665:	gap of unknown length	
* 38666	40077:	contig of 1412 bp in length	
* 40078	40177:	gap of unknown length	
* 40178	41278:	contig of 1101 bp in length	
* 41279	41378:	gap of unknown length	
* 41379	42846:	contig of 1468 bp in length	
* 42847	42946:	gap of unknown length	
* 42947	44779:	contig of 1833 bp in length	
* 44780	44879:	gap of unknown length	
* 44880	47030:	contig of 2151 bp in length	
* 47031	47130:	gap of unknown length	
* 47131	48846:	contig of 1716 bp in length	
* 48847	48946:	gap of unknown length	
* 48947	50071:	contig of 1125 bp in length	
* 50072	50171:	gap of unknown length	
* 50172	51617:	contig of 1446 bp in length	
* 51618	51717:	gap of unknown length	
* 51718	53821:	contig of 2104 bp in length	
* 53822	53921:	gap of unknown length	
* 53922	55611:	contig of 1690 bp in length	
* 55612	55711:	gap of unknown length	
* 55712	5525:	contig of 1814 bp in length	
* 5726	57625:	gap of unknown length	
* 57626	59106:	contig of 1481 bp in length	
* 59107	59206:	gap of unknown length	
* 59207	61142:	contig of 1936 bp in length	
* 61143	61242:	gap of unknown length	
* 61243	64111:	contig of 3169 bp in length	
* 64112	64511:	gap of unknown length	
* 64512	66861:	contig of 2175 bp in length	
* 66862	66786:	gap of unknown length	
* 66787	68966:	contig of 2180 bp in length	
* 68967	69066:	gap of unknown length	
* 69067	72130:	contig of 3084 bp in length	
* 72131	72230:	gap of unknown length	
* 72231	73778:	contig of 1548 bp in length	
* 73779	73878:	gap of unknown length	
* 73879	76273:	contig of 2395 bp in length	
* 76274	76373:	gap of unknown length	
* 76374	78932:	contig of 2559 bp in length	
* 78933	79032:	gap of unknown length	
* 79033	81983:	contig of 2951 bp in length	
* 81984	82083:	gap of unknown length	
* 82084	84716:	contig of 2633 bp in length	
* 84717	84816:	gap of unknown length	
* 84817	87800:	contig of 2984 bp in length	
* 87801	87900:	gap of unknown length	
* 87901	92353:	contig of 4453 bp in length	
* 92354	92453:	gap of unknown length	
* 92454	96499:	contig of 4046 bp in length	
* 96500	96599:	gap of unknown length	
* 96600	99503:	contig of 2904 bp in length	
* 99504	99603:	gap of unknown length	
* 99604	103250:	contig of 3647 bp in length	
* 103251	103350:	gap of unknown length	
* 103351	106405:	contig of 3055 bp in length	
Query Match	49.78:	Score 554; DB 2; Length 156175;	
Best Local Similarity	78.79:	Prod. No. 2.8e-116;	
Matches 735; Conservative	0;	Mismatches 170; Indels 29; Gaps 5;	
Qy 163	GATGCTTCTCCAGACTGTGGAGCTCTCTATACCAATTTGGCCCTCTCTCCAATC	222	
Db 77048	GAATACTCTTCAGATCTCTCTACACCTATACCTCTGG-----TC	77100	
Qy 223	CACAGGAAGTCCCTCTGATTCGAATGCTGCTGGAGCTCTGACCTGCCCCAAA	282	
Db 77101	CTCGGGAAGCCCTCCCATCTCTATGCTGCTGGAGCTCTGGACCTGTGCCAAA	77160	
Qy 283	TGGTGGCCCTGTGACCTGTTTCAAGACCTACCGCTTCTACTGGACCATGATGC	342	
Db 77161	TGATACTCGCCCTGGTTCAGGATCTCAAGGTTTACTGGCACCTTGGATGC	77220	

Qy 343	CTTCGTGAAGATCGTCAGGACGAGGGGACAGACCTCTGAGGGGCTCCCGGCAC	402	
Db 77221	CTTTGTGAAGATTTGTAGGC--CGAGGGACACAGGACCTGTGAGGGGCTCCAGTAC	77279	
Qy 403	CTTCGTGATGACTGTCCAGCTACCGGCATCTACTTCACTGCTATGACCAACTGMAAGC	462	
Db 77280	CTGATGATGACTCTATAGCTACCGGCATCTATTTCTTGTATGACCAACTCAAG--	77337	
Qy 463	CTTCCTGTGTGTCGAGCCCTGACCTTGACCTTCTAGCAACCCATGTTGGCTGGCGCT	522	
Db 77338	--ACCTATGTGTCAATCTTGACCTCTGACCTCTACGAAACCCAGTGTGGTGGCGT	77395	
Qy 523	GGCCGGCTGGGACCGTCACTGTCATGAGCCCCCTGGAGCTTATGCGACAAAGCTGCA	582	
Db 77396	CACCCGAATGGGACCGTGACAGTGTGAGCCCCCTGGAGCTTGTGCGAACCAAGCTGCA	77455	
Qy 583	GGCTCAGCATGTCTCTACCGGGAGCTGGGTGCTGTGTTGAACTGCACTGGCTCAGGG	642	
Db 77456	GGCTCAGCATGTCTCTACCGGAACTGGCTACCTGTGTTCAAACTGTAGCAGTTCAGAG	77515	
Qy 643	TGGTGGCTCACTGTGCTGGGCTGGGCCCCACCTGCCCCCTCGAGATGTGCCCTTCTC	702	
Db 77516	TGACTGGCGCTTCTGTGCTGGTGGGCTTCCATAGCTCTTCAAGATGCAACCTTCGC	77575	
Qy 703	AGCCCTGTACTGGTCAACTATGAGCTGTTGAAGCTGGCTCAATGGGCTCAGCCGAA	762	
Db 77576	AGCTCTCCACTGGTTCAACTCAAGCTGGTGAAGGGCCAGCTGAATGGGCCAAGACTGAA	77635	
Qy 763	GGACGACTTCTGTGGGCACTGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGC	822	
Db 77636	AGACGACACACCGTGGACATCAGTTTGTGTCTAATGGCACTCAGAAATGGGGGCTGC	77695	
Qy 823	AGTCTCACTCTACCTTTGAGCTGGTAAAGACCCACCCAGCTGGCTCTGGAGCGAT	882	
Db 77696	CACCTTTACCTTACCTTTGTGATGATGAAGACACAGACGACAAATGTCACTGTGAGCGT	77755	
Qy 883	GGAGCTGTGAGAGTGAACCCCTGCATGTGGACTGCACCTGGCTGCTCGGAGGAT	942	
Db 77756	GGAGCTATAGAG-----GTGGCTCCAGCTGGCTGGCTTCGGAGAAAT	77800	
Qy 943	CCGGCCGAGTCCGGGACCAAGGAGCTTTTTCAGAGCTTCTTCCTCGATCATCAAGGC	1002	
Db 77801	CCAGCTGATCTGGCACACAGGAGCTC--TGCAGGTTTCTTCCCAGATCATCAAGGC	77858	
Qy 1003	TGCCCTCTCTGTGCACTATGATGACACCTATGAGTTGGCAAAAGCTTCTTCCAGAG	1062	
Db 77859	TGCACCTCTCTGTGCACTATGATGAGTTGGCAAAAGCTTCTTCCAGAG	77918	
Qy 1063	GCTGAACGAGGACCGCTTCTGGGGCTGAAAG	1096	
Db 77919	GCTCAACCAAAAGCAGCTCTGGGCCACTGAAAG	77952	
RESULT 15			
AC124203	Rattus norvegicus clone CH210-75J9, *** SEQUENCING IN PROGRESS ***		
LOCUS	60 unordered pieces.		
DEFINITION	AC124203.2 GI:21431096		
ACCESSION	AC124203		
VERSION	HTG; HTGS_PHASE1.		
KEYWORDS	Rattus norvegicus.		
SOURCE	Rattus norvegicus		
ORGANISM	Bukaryota; Euthera; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 195701)		
AUTHORS	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaraturge, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,		

Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrill, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
 Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, D., Korvan, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maneshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhiney, E., McLeod, W.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pul, L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojudo, K., Rolfe, W., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sison, I.,
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 195701)
 Worley, K.C.
 Direct Submission
 Submitted (13-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 195701)
 Worley, K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 17, 2002 this sequence version replaced gi:21397217.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GYVK
 Center clone name: CH230-75J9
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 143647 bases at least Q40
 Consensus quality: 149058 bases at least Q30
 Consensus quality: 153360 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * [see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html].
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 60 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1426: contig of 1426 bp in length
 * 1427
 * 1526: gap of unknown length
 * 1527
 * 2912: contig of 1386 bp in length
 * 2913
 * 3012: gap of unknown length
 * 3013
 * 4351: contig of 1139 bp in length
 * 4352
 * 4251: gap of unknown length
 * 4252
 * 5629: contig of 1378 bp in length
 * 5630
 * 5729: gap of unknown length
 * 5730
 * 6926: contig of 1197 bp in length
 * 6927
 * 7026: gap of unknown length
 * 7027
 * 8163: contig of 1137 bp in length
 * 8164
 * 8263: gap of unknown length
 * 8264
 * 9582: contig of 1319 bp in length
 * 9583
 * 9682: gap of unknown length
 * 9683
 * 11389: contig of 1707 bp in length
 * 11390
 * 11489: gap of unknown length
 * 11490
 * 13054: contig of 1565 bp in length
 * 13055
 * 13154: gap of unknown length
 * 13155
 * 14410: contig of 1256 bp in length
 * 14411
 * 14510: gap of unknown length
 * 14511
 * 15590: contig of 1080 bp in length
 * 15591
 * 15690: gap of unknown length
 * 15691
 * 16982: contig of 1292 bp in length
 * 16983
 * 17082: gap of unknown length
 * 17083
 * 18917: contig of 1835 bp in length
 * 18918
 * 19017: gap of unknown length
 * 19018
 * 20730: contig of 1713 bp in length
 * 20731
 * 20830: gap of unknown length
 * 20831
 * 22139: contig of 1309 bp in length
 * 22140
 * 22239: gap of unknown length
 * 22240
 * 23606: contig of 1367 bp in length
 * 23607
 * 23706: gap of unknown length
 * 23707
 * 25243: contig of 1537 bp in length
 * 25244
 * 25343: gap of unknown length
 * 25344
 * 26593: contig of 1250 bp in length
 * 26594
 * 26993: gap of unknown length
 * 26994
 * 28414: contig of 1721 bp in length
 * 28415
 * 28514: gap of unknown length
 * 28515
 * 31435: contig of 2921 bp in length
 * 31436
 * 31535: gap of unknown length
 * 31536
 * 33531: contig of 1996 bp in length
 * 33532
 * 33631: gap of unknown length
 * 33632
 * 36326: contig of 2695 bp in length
 * 36327
 * 36426: gap of unknown length
 * 36427
 * 40073: contig of 3647 bp in length
 * 40074
 * 40173: gap of unknown length
 * 40174
 * 42131: contig of 1958 bp in length
 * 42132
 * 42231: gap of unknown length
 * 42232
 * 44533: contig of 2222 bp in length
 * 44534
 * 44553: gap of unknown length
 * 44554
 * 46183: contig of 1630 bp in length
 * 46184
 * 46283: gap of unknown length
 * 46284
 * 48384: contig of 2101 bp in length
 * 48385
 * 48484: gap of unknown length
 * 48485
 * 52244: contig of 3760 bp in length
 * 52245
 * 52344: gap of unknown length
 * 52345
 * 54687: contig of 2343 bp in length
 * 54688
 * 54787: gap of unknown length
 * 54788
 * 57211: contig of 2424 bp in length
 * 57212
 * 57311: gap of unknown length
 * 57312
 * 59198: contig of 1887 bp in length
 * 59199
 * 59298: gap of unknown length
 * 59299
 * 61525: contig of 2227 bp in length
 * 61526
 * 61625: gap of unknown length
 * 61626
 * 64137: contig of 2512 bp in length
 * 64138
 * 64237: gap of unknown length
 * 64238
 * 66994: contig of 2757 bp in length
 * 66995
 * 67094: gap of unknown length
 * 67095
 * 69582: contig of 2388 bp in length
 * 69583
 * 69582: gap of unknown length

	*	69583	72067:	contig of 2485 bp in length
	*	72068	72167:	gap of unknown length
	*	72168	74228:	contig of 2061 bp in length
	*	74229	74328:	gap of unknown length
	*	74329	78279:	contig of 3951 bp in length
	*	78280	78379:	gap of unknown length
	*	78380	81948:	contig of 3569 bp in length
	*	81949	82048:	gap of unknown length
	*	82049	85828:	contig of 3780 bp in length
	*	85829	90452:	contig of 4524 bp in length
	*	90453	90552:	gap of unknown length
	*	90553	93607:	contig of 3055 bp in length
	*	93608	93707:	gap of unknown length
	*	93708	96753:	contig of 3046 bp in length
	*	96754	96853:	gap of unknown length
	*	96854	99186:	contig of 2333 bp in length
	*	99187	99286:	gap of unknown length
	*	99287	103012:	contig of 3726 bp in length
	*	103013	103112:	gap of unknown length
	*	103113	107318:	contig of 4206 bp in length
	*	107319	107418:	gap of unknown length
	*	107419	112328:	contig of 4910 bp in length
	*	112329	112428:	gap of unknown length
	*	112429	117734:	contig of 5306 bp in length
	*	117735	117834:	gap of unknown length
	*	117835	121612:	contig of 3778 bp in length
	*	121613	121712:	gap of unknown length
	*	121713	125712:	contig of 4000 bp in length
	*	125713	125812:	gap of unknown length
	*	125813	130929:	contig of 5117 bp in length
	*	130930	131029:	gap of unknown length
	*	131030	135655:	contig of 4626 bp in length
	*	135656	135755:	gap of unknown length
	*	135756	140212:	contig of 4457 bp in length
	*	140213	140312:	gap of unknown length
	*	140313	143560:	contig of 3248 bp in length

Query Match 49.6%; Score 552.8; DB 2; Length 195701;

Best Local Similarity 83.4%; Pred. No. 5.1e-116;

Matches 664; Conservative 0; Mismatches 127; Indels 5; Gaps

QY	302	TGTTTCAAGACCCCTACCGCTTCCTGGCACCATGGATGCCCTTGGAAGATCGTAGG	361
Db	72766	TGTTTTCAGATCCTACCGGTTCACCTGGCACCTTGGATGCCTTGGTAAGATTGTGAGG	72825
QY	362	CACGAGGCCACAGGACCCCTCTGGAGCGGCTCCCCGCCACCCCTGGTGTGACTGTGCCA	421
D'	2826	CAGGAGGCCACAGGACCCCTGTGGAGCGGCTCCCAGCCACCCTGGTGTGACTGTACCA	72885
QY	422	GCTACCGGCATCTACTTCTACTGCTATGACAACTCAAGAGCCTTCCTGTGTGTGAGCC	481
Db	72886	GCACCTGCCATCTACTTCTACTGCTTATGACAACTCAAGAGCCTTCCTATGTGTGTGAGCC	72945
QY	482	CTGACCTCTGACCTCTACCCACCATGGTGGCGCGCTGGCCCCCGCTGGGCACCGTG	541
Db	72946	TTGACCTCTGACCTCTATGCACCAATGGTGGCTGTGTGGCCCTCGCCCAAATGGGCACCTGTG	73005
QY	542	ACTGTGATCAGCCCCCTTGAGAGCTATGCGGCACAAAGCTCAGGGCTCAGCATGTGTGCTAC	601
Db	73006	ACGGTGTGTAGCCCCCTTGAGCTTGTGTGGAACCAAGCTCAGGGCTCAGCGTGTGTCTATCC	73065
QY	602	CSGGAGCTGGGTGGCTGTGTTTGCAACTGTCAAGTGGCTCAGGTGGCTGGCGCTCACTGTGG	661
Db	73066	CCGCAACTGGCTGCCCTGTGTTTCAAGCTTGCAGTGGCTCAAGTGGGTGGTGTGCTCGCTGTGG	73125
QY	662	CTGGCTGGGGGCCCACTGCCCTTCGAGATGTGCCCTCTCAGACCCCTGTACTGTTTCAAC	721
Db	73126	CTGGGCTGGGGTCCCACTGCAAGCTCTTCGAGATGTACTCTTCTCAGCTCTTACTGTTTCACT	73185
QY	722	TATGAGCTGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAGGACCAAGCTCTGTGGGC	781
Db	73186	TATGAGCTGTGAAGAGCCAGTGAATGGGCCAAGACCCAAGAGCAGAATGTGTGGGC	73245

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:31:07 ; Search time 255.466 Seconds
(without alignments)
9820.189 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagcttcagatgctg.....aggggcaaggaggcaaggac 1114

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185219 seqs, 1125995159 residues
To: number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 101002:*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1114	100.0	1114	24	AA171019 Human mitochondria
2	1070	95.1	1662	21	AAC76992 Human ORF ORF2547
3	1056	94.8	1560	21	AAC90452 Human uncoupling p
4	1054.4	94.6	1546	24	AA171020 Human mitochondria
5	472	42.4	704	24	AB199647 Mouse ischaemic co
6	472	42.4	2170	24	ABK34735 Human cDNA for nov
7	448.4	40.3	758	24	ABL01419 Murine apoptosis r
8	242.2	21.7	5712	21	AC76691 Human ORF ORF2246
9	185.8	16.7	453	24	ABL81508 Human ovarian canc

c	10	170.2	15.3	512	24	ABK62422	Rat sequence diffe
	11	170	15.3	537	21	AAA06606	Human immunogenic
	12	170	15.3	537	22	AAS63815	Human prostate cDN
	13	170	15.3	537	22	AAH93722	Human prostate-spe
	14	170	15.3	537	22	AAH85036	Human prostate-spe
	15	170	15.3	537	22	AAH02787	Prostate tumour an
	16	170	15.3	537	24	ABL95186	Human CGI-69 cDNA
	17	154	13.8	275	22	AA100555	Human reproductive
	18	154	13.8	275	23	ABK72077	Human cDNA encodin
	19	154	13.8	275	24	ABK91669	cDNA encoding nove
	20	151.4	13.6	1258	24	ABL89865	Human polynucleoti
	21	142.6	12.8	397	21	AAH30252	Human colon cancer
	22	130.4	11.7	1449	23	ABL04893	Drosophila melanog
	23	130.4	11.7	3928	23	ABL04892	Drosophila melanog
c	24	122	11.0	422	22	AAK53810	Murine transport a
	25	112.4	10.1	1445	23	ABL24583	Drosophila melanog
c	26	93.4	8.4	452	24	ABN94635	Gene #1133 used to
	27	79.6	7.1	3645	23	ABL24582	Drosophila melanog
	28	78.2	7.0	1529	23	ABL24581	Drosophila melanog
c	29	77.4	6.9	2516	24	ABL90270	Human polynucleoti
	30	74.8	6.7	452	22	ABA59262	Human foetal liver
c	31	74.8	6.7	452	22	AAK07474	Human brain expres
	32	74.8	6.7	452	22	AAK33259	Human bone marrow
c	33	74.8	6.7	452	22	AAI39053	Probe #7739 used t
	34	74.8	6.7	452	24	ABS08091	Human genome-deriv
c	35	74.8	6.7	496	22	ABA59322	Human foetal liver
	36	74.8	6.7	496	22	AAK07540	Human brain expres
	37	74.8	6.7	496	22	AAK33336	Human bone marrow
	38	74.8	6.7	496	22	AAI39123	Probe #7809 used t
c	39	74.8	6.7	496	24	ABS08171	Human genome-deriv
	40	63.8	5.7	367	24	ABL01420	Murine apoptosis r
	41	61	5.5	997	21	AAZ61628	cDNA encoding muri
	42	61	5.5	997	22	AAQ95561	Skin cell cDNA, SE
	43	61	5.5	997	24	ABL34713	Murine cDNA isolat
	44	61	5.5	1808	22	AAQ99806	Skin cell cDNA, SE
	45	61	5.5	1808	24	ABL34958	Murine cDNA isolat

ALIGNMENTS

RESULT 1
AA171019
ID AA171019 standard; cDNA: 1114 BP.
XX AA171019;
XX AA171019;
DT 18-MAR-2002 (first entry)
XX Human mitochondrial carrier protein CGI-69L cDNA.
DE CGI-69L; mitochondrial carrier protein; human; metabolic disease;
KW obesity; cachexia; tumour; cancer; infection; immunomodulator;
KW antitumour; virucide; antibacterial; anorectic; antidiabetic;
KW brown adipose tissue; diagnosis; gene therapy; splice variant; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 14..1093
FT /*tag= a
XX
PN WC200198355-A2.
XX
PD 27-DEC-2001.
XX
PF 22-JUN-2001; 2001WO-US20117.
XX
PR 22-JUN-2000; 2000US-213307P.
XX
PA (GETH) GENENTECH INC.
PA (CURA-) CURAGEN CORP.
XX

PI Lewin D, Adams SH, Yu XX;
 XX WPI; 2002-114569/15.
 DR P-PSDB; AAM50569.
 XX
 XX CGI-69 polypeptides and polynucleotides useful for treating metabolic
 PT disorders including cachexia, obesity, diabetes and cancers -
 FT
 XX
 XX Claim 1(a); Page 9-10; 107pp; English.
 PS
 XX
 CC The present sequence is that of cDNA encoding human CGI-69L (see
 CC AAM50569), a novel splice variant of human CGI-69. Various
 CC clones were isolated from human liver upon PCR amplification and
 CC cloning. Numerous clones diverged from the previously known CGI-69
 CC sequence in that they encoded an 8-amino acid insert preceded by a
 CC W64L change, and were termed the 'long version' isoform or CGI-69L.
 CC The invention relates to a novel characterization of CGI-69 as a
 CC mitochondrial carrier protein, the discovery of CGI-69L, and the
 CC discovery of the alteration of the mitochondrial membrane potential
 CC through overexpression of carboxyl-FLAG-tagged CGI-69. CGI-69
 CC (including splice variant) nucleic acids and proteins are useful
 CC for diagnosing and treating metabolic diseases in humans, including
 CC obesity, cachexia and diabetes. Evidence for an important function
 CC for CGI-69 in modifying mitochondrial membrane potential in brown
 CC adipose tissue (BAT) is presented. The mouse orthologue of CGI-69
 CC is up-regulated in cold-treated BAT. CGI-69 may be involved in
 CC cellular thermogenic uncoupling and, therefore, may be used to
 CC diagnose and treat specific perturbations in metabolic pathways.
 CC Altering the expression of CGI-69 through gene therapy provides a
 CC means of treating metabolic diseases, such as obesity or cachexia,
 CC or of increasing or decreasing body weight. Decreasing the
 CC activity of CGI-69 can be used to treat cachexia, tumours, cancers,
 CC viral infections and bacterial infections, while increasing its
 CC activity can be used to treat obesity, tumours, cancers, viral
 CC infections and bacterial infections (all claimed). Methods are
 CC also claimed for determining whether a compound up-regulates or
 CC down-regulates expression of a CGI-69 gene, and of screening for
 CC a mutation in the CGI-69 gene.
 XX
 SQ Sequence 1114 BP; 194 A; 349 C; 332 G; 239 T; 0 other;

Query Match 100.0%; Score 1114; DB 24; Length 1114;
 Best Local Similarity 100.0%; Pred. No. 4.3e-264;
 Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAAGCTTCAAGATGGCTGACGAGGACCTGCGGGGATCAGCCCTCCAGCAATGGT 60
 DB 1 CTGAAGCTTCAAGATGGCTGACGAGGACCTGCGGGGATCAGCCCTCCAGCAATGGT 60
 QY 61 GGCCTCAGGACCGGGGCTGTGGTTACTCTCTTCTATGACACCCCTGGAGTGTGAA 120
 DB 61 GGCCTCAGGACCGGGGCTGTGGTTACTCTCTTCTATGACACCCCTGGAGTGTGAA 120
 QY 121 GGTTCGCTGAGTCTCAGGGGCTTCCATGCGGAGGATGATGCTTCTCCAGACT 180
 DB 121 GGTTCGCTGAGTCTCAGGGGCTTCCATGCGGAGGATGATGCTTCTCCAGACT 180
 QY 181 GTGGAGCTCTCTATACCAATTTGCCCTCTCTCTCCAAATCCAGGGAAGTCTCT 240
 DB 181 GTGGAGCTCTCTATACCAATTTGCCCTCTCTCTCCAAATCCAGGGAAGTCTCT 240
 QY 241 GTATTGCAATGGTGTCTGAGGCTCTGTACTGTGCGCAATGGTGGCGGTGGCCAC 300
 DB 241 GTATTGCAATGGTGTCTGAGGCTCTGTACTGTGCGCAATGGTGGCGGTGGCCAC 300
 QY 301 CTGGTTTCAGACCTACCGGCTTCACTGCGACCATGATGCTCTCTGGAAGATCTGAG 360
 DB 301 CTGGTTTCAGACCTACCGGCTTCACTGCGACCATGATGCTCTCTGGAAGATCTGAG 360
 QY 361 GCACGAGGACACGAGGACCTCTGGAGCGGCTCTCCCGCCACCTGGTGTGATGTGCC 420
 DB 361 GCACGAGGACACGAGGACCTCTGGAGCGGCTCTCCCGCCACCTGGTGTGATGTGCC 420

QY 421 AGCTACGCCCATCTACTTCACTGCTATGACCAACTCAAGGCTTCTGTGTGTCGAGC 480
 DB 421 AGCTACGCCCATCTACTTCACTGCTATGACCAACTCAAGGCTTCTGTGTGTCGAGC 480
 QY 481 CCGTACCTCTGACTCTAGCACCACCATGGTGGCTGGCGCGCTGGCCCGCTGGGCAACCGT 540
 DB 481 CCGTACCTCTGACTCTAGCACCACCATGGTGGCTGGCGCGCTGGCCCGCTGGGCAACCGT 540
 QY 541 GACTGTGATCAGCCCTCTGAGCTTTATCGCGACAAAGCTGCGAGGCTCAGCATGTGTGTA 600
 DB 541 GACTGTGATCAGCCCTCTGAGCTTTATCGCGACAAAGCTGCGAGGCTCAGCATGTGTGTA 600
 QY 601 CCGGAGCTGGGTGGCTGTGTTGCAACTGAGTGGCTCAGGTTGGTGGCTGCTCACTGTG 660
 DB 601 CCGGAGCTGGGTGGCTGTGTTGCAACTGAGTGGCTCAGGTTGGTGGCTGCTCACTGTG 660
 QY 661 GCTGGGCTGGGGGCGCCCACTGCTTTCGAGATGTGCTTCTCAGCCCTGTACTGTGTTCAA 720
 DB 661 GCTGGGCTGGGGGCGCCCACTGCTTTCGAGATGTGCTTCTCAGCCCTGTACTGTGTTCAA 720
 QY 721 CTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCCGAGGACAGACTTCTGTGGG 780
 DB 721 CTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCCGAGGACAGACTTCTGTGGG 780
 QY 781 CATGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGAGTGTGCTTACCTT 840
 DB 781 CATGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGAGTGTGCTTACCTT 840
 QY 841 TGACGTGTGTAAGACCCCAACCGCAGCTGCTCTGGGAGCGATGAGGCTGTGAGAGTGA 900
 DB 841 TGACGTGTGTAAGACCCCAACCGCAGCTGCTCTGGGAGCGATGAGGCTGTGAGAGTGA 900
 QY 901 CCCCTGATGTGAGCTTCACTGGCTGCTGCTGCGGAGGATCGGGCCGAGTGGGCGAC 960
 DB 901 CCCCTGATGTGAGCTTCACTGGCTGCTGCTGCGGAGGATCGGGCCGAGTGGGCGAC 960
 QY 961 CAAGGAGCTCTTTTCAGGCTTCTCTTCGGATCATCAAGGCTGCCCCCTCTGTGTCAT 1020
 DB 961 CAAGGAGCTCTTTTCAGGCTTCTCTTCGGATCATCAAGGCTGCCCCCTCTGTGTCAT 1020
 QY 1021 CATGATCAGCACTTATGAGTTGGCAAAAGCTTCTTCCAGAGCTGAACAGGACCGCT 1080
 DB 1021 CATGATCAGCACTTATGAGTTGGCAAAAGCTTCTTCCAGAGCTGAACAGGACCGCT 1080
 QY 1081 TCTGGGCGGCTGAAAGGGCGGAGGAGGCAAGGAC 1114
 DB 1081 TCTGGGCGGCTGAAAGGGCGGAGGAGGCAAGGAC 1114
 RESULT 2
 AAC76992
 ID AAC76992 standard; cDNA; 1662 BP.
 XX
 AC AAC76992;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2547 polynucleotide sequence SEQ ID NO:5093.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoiatic; antiparkinsonian; noctropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.
OS
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR
XX
PT novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 5; Page 4281-4282; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in BAB40237 to BAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antiproliferative; antiparkinsonian; neurotropic; immunoprotective;
CC osteoprotective; anticonvulsant; antiarthritic; immunosuppressive;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy.
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
S_q Sequence 1662 BP; 336 A; 516 C; 468 G; 341 T; 1 other;
Query Match 96.1%; Score 1070; DB 21; Length 1662;
Best Local Similarity 97.5%; Pred. No. 3.2e-253;
Matches 1104; Conservative 0; Mismatches 10; Indels 18; Gaps 1;
QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCCCTGGGGCATCAGCCCCCTCCAGCAAAATGGT 60
DB 123 CTGAAGCTTCAAGATGGCTGACAGGACCCCTGGGGCATCAGCCCCCTCCAGCAAAATGGT 182
QY 61 GGCTCAGGACCCGGGGTGGTGTACCTCTCTTCATGACACCCCTGACCGTGGTAA 120
DB 183 GGCTCAGGACCCGGGGTGGTGTACCTCTCTTCATGACACCCCTGACCGTGGTAA 242
QY 121 GGTTCGCTCGAGTCTCAGCGGCCCTCCATGGCCAGCGAGTGTCTTCCTCCAGACT 180
DB 243 GGTTCGCTCGAGTCTCAGCGGCCCTCCATGGCCAGCGAGTGTCTTCCTCCAGACT 302
QY 181 GTGAGCCCTCTCTATACCAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTGCCTCT 240
DB 303 GTGAGCCCTCTCTATACCAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTGCCTCT 362
QY 241 GTATTGCAATGGTCTCTGGAGCCCTCTGTACCTGTGGCCCAATGGTGCCTGTGCCAC 300

DB 363 GTATTGCAATGGTCTCTGGAGCCTCTGTACCTGTGCCCAAAATGGTGCCTGTGTGCCAC 422
QY 301 CTGTTTCAAGACCCCTACCCGCTTCACTGCGCACCATTGATGCCTTCTGTAAGATCTGTGAG 360
DB 423 CTGTTTCAAGACCCCTACCCGCTTCACTGCGCACCATTGATGCCTTCTGTAAGATCTGTGAG 482
QY 361 GCACGAGGACACACGAGCCCTCTGAGACCGCCCTCCCGCCACCCCTGGTGTGATGACTGTGCC 420
DB 483 GCACGAGGACACACGAGCCCTCTGAGACCGCCCTCCCGCCACCCCTGGTGTGATGACTGTGCC 542
QY 421 AGTACCGGCATCTACTTCACTGCTATGACCAACTGAAAGGCTTCTCTGTGTGATGAGC 480
DB 543 AGTACCGGCATCTACTTCACTGCTATGACCAACTGAAAGGCTTCTCTGTGTGATGAGC 602
QY 481 CCTGACCTCTGACCTCTAGCACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 603 CCTGACCTCTGACCTCTAGCACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
QY 541 GACTGTGATCAGCCCTCTGAGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGTGTA 600
DB 663 GACTGTGATCAGCCCTCTGAGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGTGTA 722
QY 601 CCGGAGCTGGGTGCTGCTGTGTAAGTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 660
DB 723 CCGGAGCTGGGTGCTGCTGTGTAAGTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 782
QY 661 GCTGGCTGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
DB 783 GCTGGCTGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
QY 703 AGCCCTGTACTGTTTCAACTATGAGCTGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 762
DB 843 AGCCCTGTACTGTTTCAACTATGAGCTGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 902
QY 763 GACACAGACTTCTGTGGGCTGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
DB 903 GACACAGACTTCTGTGGGCTGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
QY 823 AGTGCTGACTCTACCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
DB 963 AGTGCTGACTCTACCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
QY 883 GGAGGCTGTGAGAGTGAACCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942
DB 1023 GGAGGCTGTGAGAGTGAACCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
QY 943 CCGGGCCGAGTGGGCGACCAAGGAGCTTCTTGGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1002
DB 1083 CCGGGCCGAGTGGGCGACCAAGGAGCTTCTTGGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
QY 1003 TGCCCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
DB 1143 TGCCCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
QY 1063 GCTGAACACGAGCCGCTTCTGCGGCGCTGAAAGGGGCAAGGGGCAAGGAGGAC 1114
DB 1203 GCTGAACACGAGCCGCTTCTGCGGCGCTGAAAGGGGCAAGGGGCAAGGAGGAC 1254
RESULT 3
AAC90452
ID AAC90452 standard; cDNA; 1560 BP.
XX
AC AAC90452;
XX
DT 12-MAR-2001 (first entry)
XX
DE Human uncoupling protein cDNA #1.
XX
KW Human; uncoupling protein; immunosuppressive; antiarthritic;
KW antineumatic; antiproliferative; cardiant; vasotropic;
KW cerebroprotective; neuroprotective; antibacterial; ophthalmological;
KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

xx PF 22-JUN-2001; 2001WO-US20117.
xx PR 22-JUN-2000; 2000US-213307P.
xx (GETH) GENENTECH INC.
xx PA (CURA-) CURAGEN CORP.
xx PL Lewin D, Adams SH, Yu XX;
xx DR WPI; 2002-114569/15.
xx CGI-69 polypeptides and polynucleotides useful for treating metabolic
xx disorders including cachexia, obesity, diabetes and cancers -
xx Disclosure; Page 11-12; 107pp; English.
xx The present sequence is that of cDNA encoding human CGI-69 (see
xx AM50570). Various CGI-69 clones were isolated from human liver
xx upon PCR amplification and cloning. Some diverged from the
xx previously known CGI-69 sequence in that they encoded an 8-amino
xx acid insert preceded by a W6L change, and were termed the 'long
xx version' isoform or CGI-69L (see AAI71019). The invention relates
xx to a novel characterization of CGI-69 as a mitochondrial carrier
xx protein, the discovery of CGI-69L, and the discovery of the
xx alteration of the mitochondrial membrane potential through
xx overexpression of carboxyl-FLAG-tagged CGI-69. CGI-69 (including
xx splice variant) nucleic acids and proteins are useful for
xx diagnosing and treating metabolic diseases in humans, including
xx obesity, cachexia and diabetes. Evidence for an important function
xx for CGI-69 in modifying mitochondrial membrane potential in brown
xx adipose tissue (BAT) is presented. The mouse orthologue of CGI-69
xx is up-regulated in cold-treated BAT. CGI-69 may be involved in
xx cellular thermogenic uncoupling and, therefore, may be used to
xx diagnose and treat specific perturbations in metabolic pathways.
xx Altering the expression of CGI-69 through gene therapy provides a
xx means of treating metabolic diseases, such as obesity or cachexia,
xx or of increasing or decreasing body weight. Decreasing the
xx activity of CGI-69 can be used to treat cachexia, tumours, cancers,
xx viral infections and bacterial infections, while increasing its
xx activity can be used to treat obesity, tumours, cancers, viral
xx infections and bacterial infections (all claimed). Methods are
xx also claimed for determining whether a compound up-regulates or
xx down-regulates expression of a CGI-69 gene, and of screening for
xx a mutation in the CGI-69 gene.
xx SQ Sequence 1546 BP; 272 A; 489 C; 459 G; 326 T; 0 other;

Db 381 CTGGTTTCAAGACCCCTACCCGCTTCTCTGTCACCATGATGCTTCTGTAAGATCGTGAG 440
Qy 361 GCACGAGGGGACACGAGACCTCTGAGCGGCCTCCCGCCACCTGTGTGATGACTGTGCC 420
Db 441 GCACGAGGGGACACGAGACCTCTGAGCGGCCTCCCGCCACCTGTGTGATGACTGTGCC 500
Qy 421 AGCTACCGGCATCTACTTCTACTGCTATGACCACTGAAGGCTTCTCTGTGTGTGAGC 480
Db 501 AGCTACCGGCATCTACTTCTACTGCTATGACCACTGAAGGCTTCTCTGTGTGTGAGC 560
Qy 481 CCTGACCTCTGACCTCTACGACCCCATGCTGCGCGCTGCGCCGCTTGGSCACCGT 540
Db 561 CCTGACCTCTGACCTCTACGACCCCATGCTGCGCGCTGCGCCGCTTGGSCACCGT 620
Qy 541 GACTGTGATCAGCCCTCTGAGCTTATCGGACAAAGCTGAGGCTCAGCATGTGTGTA 600
Db 621 GACTGTGATCAGCCCTCTGAGCTTATCGGACAAAGCTGAGGCTCAGCATGTGTGTA 680
Qy 601 CCGGAGCTGGGTGCTGTGTTCGAACTGCAGTGGCTCAGGCTGGCTGGCGCTCAGCTGTG 660
Db 681 CCGGAGCTGGGTGCTGTGTTCGAACTGCAGTGGCTCAGGCTGGCTGGCGCTCAGCTGTG 740
Qy 661 GCTGGGCTGGGCGCCCACTGCTTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAA 720
Db 741 GCTGGGCTGGGCGCCCACTGCTTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAA 800
Qy 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAGGACAGACTTCTGTGG 780
Db 801 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAGGACAGACTTCTGTGG 860
Qy 781 CATGAGCTTGTGGCTGGCTGCTCATCTCAGGACGGTGGCTGCAGTGGCTGACTCTACCCCT 840
Db 861 CATGAGCTTGTGGCTGGCTGCTCATCTCAGGACGGTGGCTGCAGTGGCTGACTCTACCCCT 920
Qy 841 TGACGTGTGAAGACCCCAAGCCAGGCTGGCTCTGGAGCGATGAGGCTGTGAGAGTGAA 900
Db 921 TGACGTGTGAAGACCCCAAGCCAGGCTGGCTCTGGAGCGATGAGGCTGTGAGAGTGAA 980
Qy 901 CCCCCTGCATGTGGAATCAGCTGGCTGCTGCGGAGGATCCGGGCCGAGTCGGGCAC 960
Db 981 CCCCCTGCATGTGGAATCAGCTGGCTGCTGCGGAGGATCCGGGCCGAGTCGGGCAC 1040
Qy 961 CAAGGAGCTCTTTCGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTCTGTGCCAT 1020
Db 1041 CAAGGAGCTCTTTCGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTCTGTGCCAT 1100
Qy 1021 CATGATCAGCACTATGAGTTGGCAAAAGCTTCTTTCAGAGGCTGAACACGAGACCGGCT 1080
Db 1101 CATGATCAGCACTATGAGTTGGCAAAAGCTTCTTTCAGAGGCTGAACACGAGACCGGCT 1160
Qy 1081 TCTGGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 1114
Db 1161 TCTGGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 1194

RESULT 5
ABI99647
ID ABI99647 standard; cDNA; 704 BP.
XX AC ABI99647;
XX DT 07-MAR-2002 (first entry)
XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:675.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX OS Mus musculus.
XX PN W0200188188-A2.
XX

PD 22-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-JP04192.
 XX
 PR 18-MAY-2000; 2000JP-0145977.
 XX
 PA (UNIV-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX WPI; 2002-034733/04.
 XX
 DR
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes -
 XX
 PS Claim 2; Page 1679-1680; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding
 CC the protein sequences in AB57020 to AB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.
 XX
 SQ Sequence 704 BP; 132 A; 204 C; 209 G; 159 T; 0 other;

Query Match 42.4%; Score 472; DB 24; Length 704;
 Best Local Similarity 83.2%; Pred. No. 2.3e-106;
 Matches 561; Conservative 0; Mismatches 110; Indels 3; Caps 2;

QY 432 TCTACTTCAGCTGCTATGACCACTGAAGGCTTCTCTGTGTGCTGAGCCCTGACCTCTG 491
 DB 3 TCGATTCATGGCTTACGACCACTCAAGGCTTCTCTGTGTGCTGAGCTTGTGACCTCTG 62
 QY 492 ACCTCTACGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 551
 DB 63 ACCTCTACGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 QY 552 GCGGCTGAGCTTATGCGGACCAAGCTGCGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 611
 DB 121 GCGGCTTGGAGCTGCTGCGGACCAAGCTGCGAGCTGAGCTGCTGCTGCTGCTGCTGCTG 180
 QY 612 GTGCTGTGTTTGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 671
 DB 181 CTCTCTGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 QY 672 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
 DB 241 GTCCACAGCTCTTCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 QY 732 TGAAGCTGCTGCTCAATGGGCTCAGCGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
 DB 301 TGAAGCTGCTGCTCAGCGACTGAGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 QY 792 TGGCTGTGCTGCTCAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
 DB 361 TGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419
 QY 852 AGACCCAAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911
 DB 420 AGACACACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479
 QY 912 TGGACTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971

DB 480 TGGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
 QY 972 TTGCAGGCTTCTCTCGGATCATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1031
 DB 540 TTGCAGGCTTCTCTCGGATCATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
 QY 1032 CTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
 DB 600 CTTACGAGTGTGGCAAAAGCTTCTTCCAGAGGCTCAACAGGAGGAGGCTCTTGGCGGCT 659
 QY 1092 GAAAGGGGCAAGCA 1105
 DB 660 TGAAGTGGGAGCA 673

RESULT 6
 ABK34735
 ID ABK34735 standard; cDNA; 2170 BP.
 XX
 AC ABK34735;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA for novel secreted protein, SEQ ID 504.
 XX
 KW Human; ss; gene; secreted protein; immune deficiency; viral infection;
 KW bacterial infection; fungal infection; autoimmune disorder; burn;
 KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
 KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
 KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
 KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
 KW lymphoid cell deficiency.
 XX
 OS Homo sapiens.
 XX
 PN WO200177290-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US10295.
 XX
 PR 06-APR-2000; 2000US-194941P.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 Gulukota K, Graham JR;
 XX WPI; 2002-179323/23.
 XX
 PT Six hundred and twenty five polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for
 PT treating immune deficiencies and disorders such as autoimmune disorders
 PT -
 XX
 PS Claim 1; Page 269-270; 339pp; English.
 XX
 CC The invention relates to 625 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins, their complements and sequences that hybridise to them.
 CC Also included are a vector comprising the polynucleotide, a host cell
 CC transformed with the vector, the proteins encoded by the
 CC polynucleotides, antibodies that bind to the proteins and identification
 CC of modulators of the proteins or the expression of the polynucleotide.
 CC The polynucleotides can be used as probes for the identification
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 CC and proteins can also be used as nutritional supplements. The protein
 CC is useful in the treatment of various immune deficiencies and disorders
 CC such as viral infections, bacterial infections, fungal infections,
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 CC autoimmune thyroiditis and diabetes) and allergic reactions and

QY 1062 GGCTGAA 1068
 Db 2731 AACAAAA 2725

RESULT 9

ABL81508
 ID ABL81508 standard; cDNA; 453 BP.

AC ABL81508;
 XX 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:4486.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 KW Homo sapiens.
 OS WO200192581-A2.

PN 06-DEC-2001.

PL 29-MAY-2001; 2001WO-US17756.

PF 26-MAY-2000; 2000US-207484P.

PR (CORI-) CORIXA CORP.

PA Algate PA, Harlocker SL, Jones R;

PI WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising

PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding

PT polypeptide, antibody specific to polypeptide or T cell expressing

PT polypeptide

PS Claim 1; SEQ ID 4486; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers

CC and immunostimulants; and a polypeptide (II), of a ovarian tumour

CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence

CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to

CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell

CC population of (II), or antigen presenting cells that express (II).

CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to

CC (S1) can be used for detecting ovarian cancer in a patient's biological

CC sample preferably serum or ovarian tissue. The method comprises

CC contacting a biological sample from a patient with (IV), detecting the

CC amount of polynucleotide hybridising to (IV) and comparing the amount to

CC a predetermined cutoff value and thereby detecting ovarian cancer in the

CC patient, where the amount of polynucleotide hybridising to (IV) is

CC detected preferably by polymerase chain reaction (PCR). (I) comprising

CC (III) and/or (II) is useful for stimulating and/or expanding T cells

CC specific for an ovarian tumour protein comprising contacting T cells

CC with (III) or (II). (III) is useful in design and preparation of

CC ribozyme molecules for inhibiting expression of the tumour polypeptides

CC and proteins in tumour cells; and to isolate a full length gene from a

CC suitable library e.g., a tumour cDNA library using well known

CC techniques.

XX Sequence 453 BP; 74 A; 154 C; 128 G; 97 T; 0 other;

QY Query Match 16.7%; Score 185.8; DB 24; Length 453;

XX Best Local Similarity 92.0%; Pred. No. 4.5e-36;

XX Matches 207; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 503 CCATGTCGCTGGCCGCTGCGCCGCTGGGACCGTGTGATCAGCCCGCTGGAG 562

Db 230 CCATGTCACCCCAATCCCAACACAGTGGGACCGTGTGATCAGCCCGCTGGAG 289

QY 563 CTTATCGGACAAAGCTGAGGCTCAGCATGTGTCACCGGAGCTGGGTCGTGTT 622

Db 290 CTTATCGGACAAAGCTGAGGCTCAGCATGTGTCGTACCGGAGCTGGTGCCTGTGTT 349

QY 623 CGAACTGCACTGGCTCAGGGTGGCTGGCGCTCACTGGCTGGCTGGGCCCCCACTGCC 682

Db 350 CGAACTGCACTGGCTCAGGGTGGCTGGCGCTCACTGGCTGGCTGGGCCCCCACTGCC 408

QY 683 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 727

Db 409 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 453

RESULT 10

ABK62422/c

ID ABK62422 standard; cDNA; 512 BP.

XX AC ABK62422;

XX 18-JUN-2002 (first entry)

DT Rat sequence differentially expressed in response to a hepatotoxin #329.

DE Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;

XX differential expression; centrilobular necrosis; steatosis.

XX Rattus norvegicus.

XX WO200210453-A2.

XX 07-FEB-2002.

XX 30-JUL-2001; 2001WO-US23872.

XX 31-JUL-2000; 2000US-222040P.

XX 02-NOV-2000; 2000US-244880P.

XX 11-MAY-2001; 2001US-290029P.

XX 15-MAY-2001; 2001US-290645P.

XX 22-MAY-2001; 2001US-292336P.

XX 06-JUN-2001; 2001US-295798P.

XX 13-JUN-2001; 2001US-297457P.

XX 19-JUN-2001; 2001US-298884P.

XX 09-JUL-2001; 2001US-303459P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;

XX WPI; 2002-241625/29.

XX Predicting toxic effects of compounds or the progression of these toxic

XX effects by determining the changes in gene expression in tissues or

XX cells exposed to the toxin and comparing these to gene expression in

XX unexposed tissues or cells

XX Claim 1; Seq ID No 329; 239pp; English.

XX The invention relates to methods for predicting toxic effects of

XX compounds or the progression of these toxic effects by determining the

XX global changes in gene expression in tissues or cells exposed to the

XX toxin and comparing these to gene expression in unexposed tissues or

XX cells. Also included are methods of predicting at least one toxic

XX effect of a compound or progression of a toxic effect, preferably the

XX hepatotoxicity of a compound, comprising detecting the level of

XX expression in a tissue or cell sample exposed to the compound of two or

XX more genes listed in the specification, where differential expression of

XX the genes is indicative of at least one toxic effect or progression.

XX The method can also be used to identify an agent which modulates the

XX toxic response and predict cellular pathways that a compound modulates

XX in a cell. The methods utilise a set of at least two probes (on a solid

XX support in kit form), where each of the probes comprises a sequence that

XX specifically hybridises to a gene listed in the specification, a computer

XX system comprising a database containing information identifying the

XX expression level in a tissue or cell sample exposed to a hepatotoxin of a

CC set of genes comprising at least two genes listed in the specification.
 CC and a user interface to view the information used to present information.
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.
 XX
 SQ Sequence 512 BP; 119 A; 129 C; 147 G; 116 T; 1 other;

Query Match 15.3%; Score 170.2; DB 24; Length 512;
 Best Local Similarity 81.7%; Pred. No. 3.1e-32;
 Matches 196; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 857 CAAGCCAGGTCGCTCTGGAGCAGATGAGGCTGTGAGAGTGAACCCCTGCATGTGGAC 916

Db 511 CAGCGCAAAATGTCACTGGGAGCAGTGGAGGCTATGAGAGTGAAGCCGCAAGAGTTGAC 452

QY 917 TCCACCTGGCTGCTGCTGGAGGATCGGCGGAGTCCGGCACCAGGAGACTCTTTGCA 976

Db 451 TCCACCTGGCTGCTGCTGGAGGATCCAGGCTGAATCTGGCACCAGGAGACTCTTTGCA 392

QY 977 GGCTTCCTTCCTCGGATCATCAAGGCTGCCCTCTCTGTGCCATCATGATCAGCACCTAT 1036

Db 391 GGTTCCTCCNAGGATCATCAAGGCTGCCCTCTCTGTGCCATCATGATCAGCACCTAT 332

QY 1037 GAGTTGGGAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCTTCTGGCGCTGAAG 1096

Db 331 GAGTTGGGAAAAGCTTCTTCACAGGCTCAACGAGGACGAGCTCTGGCGCTGAAG 272

RESULT 11

AA06606
 ID AA06606 standard; cDNA; 537 BP.

XX
 AC AA06606;

XX
 DT 13-JUN-2000 (first entry)

XX
 DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:387.

XX
 KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;

XX
 KW immunogenic; cytostatic; vaccine; ss.

XX
 OS Homo sapiens.

XX
 PN WO200004149-A2.

XX
 PD 27-JAN-2000.

XX
 PF 14-JUL-1999; 99WO-US15838.

XX
 PR 14-JUL-1998; 98US-0115453.

XX
 PR 14-JUL-1998; 98US-0116134.

XX
 PR 23-SEP-1998; 98US-0159812.

XX
 PR 23-SEP-1998; 98US-0159822.

XX
 PR 15-JAN-1999; 99US-0232149.

XX
 PR 15-JAN-1999; 99US-0232880.

XX
 PR 09-APR-1999; 99US-0288946.

XX
 PA (CORI-) CORIXA CORP.

XX
 PA Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;

XX
 PI WPI; 2000-171268/15.

XX
 PT New polypeptide useful for treating and diagnosing prostate cancer

PT comprises an immunogenic portion of prostate tumor protein -
 XX
 PS Claim 50; Page 235-236; 263pp; English.

XX
 CC The present invention describes isolated polypeptides, comprising an
 CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
 CC and polynucleotides encoding them have cytostatic activity and can be
 CC used in vaccines and in gene therapy. The polypeptides and
 CC polynucleotides encoding them, antigen presenting cells which express
 CC the polypeptides, antibodies against the polypeptides and vaccines
 CC comprising them can be used for inhibiting the development of prostate
 CC cancer in a patient. The polypeptides can be used to generate antibodies
 CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
 CC the polynucleotides encoding the polypeptides can be used as a probe or
 CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
 CC AA082000 to AA082020 represent sequences used in the exemplification of
 CC the present invention.

XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 15.3%; Score 170; DB 21; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCGGAGTCGGCACCAGGAGCTCTTTCAGGCTTCTTCTCGATCATCAGGCTG 1004

Db 1 GGGCGGAGTCGGCACCAGGAGCTCTTTCAGGCTTCTTCTCGATCATCAGGCTG 60

QY 1005 CCCCTCTCTGTGCATCATGATCAGACCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGC 1064

Db 61 CCCCTCTCTGTGCATCATGATCAGACCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGC 120

QY 1065 TGAACCAAGACCGGCTTCTGGCGCTGAAAGGGGCAAGGAGGCAAGGAC 1114

Db 121 TGAACCAAGACCGGCTTCTGGCGCTGAAAGGGGCAAGGAGGCAAGGAC 170

RESULT 12

AA063815
 ID AA063815 standard; cDNA; 537 BP.

XX
 AC AA063815;

XX
 DT 29-JAN-2002 (first entry)

XX
 DE Human prostate cDNA sequence #357.

XX
 KW Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.

XX
 KW Homo sapiens.

XX
 PN WO200173032-A2.

XX
 PD 04-OCT-2001.

XX
 PF 27-MAR-2001; 2001WO-US09919.

XX
 PR 27-MAR-2000; 2000US-0536857.

XX
 PR 09-MAY-2000; 2000US-0568100.

XX
 PR 12-MAY-2000; 2000US-0570737.

XX
 PR 13-JUN-2000; 2000US-0593793.

XX
 PR 27-JUN-2000; 2000US-0605783.

XX
 PR 10-AUG-2000; 2000US-0636215.

XX
 PR 29-AUG-2000; 2000US-0651236.

XX
 PR 06-SEP-2000; 2000US-0657279.

XX
 PR 02-OCT-2000; 2000US-0679426.

XX
 PR 10-OCT-2000; 2000US-0685166.

XX
 PA (CORI-) CORIXA CORP.

XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2001-639232/73.
XX
XX New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer -
XX
XX Claim 1; Page 361; 579pp; English.
XX
XX The invention relates to isolated prostate-specific
CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC antibodies raised against the polypeptides (or antigenic epitopes
CC derived from them) and antigen-presenting cells expressing the
CC polypeptides. The antibodies are useful for detecting the presence of
CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC the antigen-presenting cells are useful for stimulating and/or expanding
CC T cells specific for a tumour protein, and for inhibiting the development
CC of cancer especially prostate cancer. Compositions comprising the
CC polynucleotide and/or polypeptide are useful for stimulating an immune
CC response, and for treating cancer. The oligonucleotide is useful for
CC detecting cancer. The present sequence is a prostate specific
CC polynucleotide of the invention.
XX
XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
SQ
Query Match 15.3%; Score 170; DB 22; Length 537;
Best Local Similarity 100.0%; Pred. No. 3.6e-32;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 945 GGGCCGAGTCGGCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGCGATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCTCTCTGCGATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 120
QY 1065 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCAAGAGGCAAGGAC 1114
Db 121 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCAAGAGGCAAGGAC 170
RESULT 13
AAH93722
ID AAH93722 standard; cDNA; 537 BP.
AC AAH93722;
DT 04-OCT-2001 (first entry)
DE Human prostate-specific cDNA sequence CGI-69.
KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
XW cytostatic; gene therapy; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
PD 19-JUL-2001.
PF 16-JAN-2001; 2001WO-US01574.
XX
PR 14-JAN-2000; 2000US-0483672.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX
XX WPI; 2001-425873/45.
DR
XX New polynucleotide encoding a prostate-specific protein, for

PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX
XX Claim 1; Page 359-360; 543pp; English.
PS
XX The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93957 to AAH93944 and AAH01115 to
CC AAH01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
SQ
Query Match 15.3%; Score 170; DB 22; Length 537;
Best Local Similarity 100.0%; Pred. No. 3.6e-32;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 945 GGGCCGAGTCGGCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGCGATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCTCTCTGCGATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 120
QY 1065 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCAAGAGGCAAGGAC 1114
Db 121 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCAAGAGGCAAGGAC 170
RESULT 14
AAH85036
ID AAH85036 standard; cDNA; 537 BP.
AC AAH85036;
DT 25-SEP-2001 (first entry)
DE Human prostate-specific cDNA sequence CGI-69.
KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
XW chromosome 22q11.2; prostate-specific protein; chromosome 1;
XW prostate specific antigen; PSA; ss.
XX
OS Homo sapiens.
XX
PN WO200134802-A2.
PD 17-MAY-2001.
PF 09-NOV-2000; 2000WO-US00904.
XX
PR 12-NOV-1999; 99US-0439313.
PR 18-NOV-1999; 99US-0443686.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, wang A;
XX
XX WPI; 2001-308785/32.
DR
XX Isolated polypeptide comprising at least an immunogenic portion of a
PT prostate-specific protein, useful in the diagnosis and therapy of

PT prostate cancer -
 XX Claim 31; Page 259; 325pp; English.
 CC The present invention describes an isolated polypeptide (P1) comprising
 CC at least an immunogenic portion of a prostate-specific protein, or its
 CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
 CC (N1) have cytostatic activity and can be used in vaccine production.
 CC The polypeptides, nucleic acids and antibodies from the present
 CC invention are useful in the diagnosis and therapy of prostate cancer.
 CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
 CC region. Prostate specific antigen (PSA) P501S was located on
 CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
 CC polynucleotide and polypeptide sequences used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 15.3%; Score 170; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 945 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCGGATCATCAAGGCTG 1004
 DB 1 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCGGATCATCAAGGCTG 60
 QY 1005 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCCAGAGGC 1064
 DB 61 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCCAGAGGC 120
 QY 1065 TGAACCAAGACCGGCTTCTGGGGCGCTGAAAGGGCAAGGAGGCAAGGAC 1114
 DB 121 TGAACCAAGACCGGCTTCTGGGGCGCTGAAAGGGCAAGGAGGCAAGGAC 170

RESULT 15
 AAH02787
 ID AAH02787 standard; cDNA; 537 BP.
 XX
 AC AAH02787;
 DT 14-JUN-2001 (first entry)
 XX
 DE Prostate tumour antigen cDNA sequence for CGI-69.
 XX
 KW Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
 KW prostate cancer; immunogenic; cytostatic; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200125272-A2.
 PD 12-APR-2001.
 XX
 PF 04-CT-2000; 2000WO-US27464.
 XX
 PR 04-OCT-1999; 99US-0157455.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Skelky YAW, Reed SG, Cheever MA;
 XX
 DR WPI; 2001-245062/25.
 XX
 PT Prostate specific protein and its encoding polynucleotide, useful for
 PT the treatment and diagnosis of prostate cancer -
 XX
 PS Claim 50; Page 244; 276pp; English.
 XX
 CC The present invention describes an isolated polypeptide (I) comprising
 CC at least an immunogenic portion of a prostate tumour antigen protein or
 CC its variant. (I) have cytostatic activity and can be used in vaccine

CC production. (I), prostate tumour antigen polynucleotides, an antigen
 CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
 CC pharmaceutical composition containing (I) are useful for inhibiting the
 CC development of cancer in a patient. Antibodies specific for prostate
 CC specific proteins and oligonucleotides that hybridise to a
 CC polynucleotide that encodes a prostate specific protein are useful
 CC for detecting the presence or absence of a cancer or monitoring the
 CC progression the progression of a cancer, especially prostate cancer.
 CC AAH02422 to AAH2872, AAH74798 to AAH74821 and AAH74830 are sequences
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 15.3%; Score 170; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 945 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCGGATCATCAAGGCTG 1004
 DB 1 GGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCGGATCATCAAGGCTG 60
 QY 1005 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCCAGAGGC 1064
 DB 61 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCCAGAGGC 120
 QY 1065 TGAACCAAGACCGGCTTCTGGGGCGCTGAAAGGGCAAGGAGGCAAGGAC 1114
 DB 121 TGAACCAAGACCGGCTTCTGGGGCGCTGAAAGGGCAAGGAGGCAAGGAC 170
 Search completed: January 22, 2003, 19:03:32
 Job time : 280.466 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:32:11 ; Search time 45.6489 seconds
(without alignments)
7484.038 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagcttcaagatgctg.....aggggcaaggaggaagac 1114

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

T. number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1054.4	94.6	1643	2	US-08-933-750C-68
2	1054.4	94.6	1643	3	US-09-234-613-68
3	170	15.3	537	4	US-09-605-785-387
4	170	15.3	537	4	US-09-439-313-387
5	170	15.3	537	4	US-09-352-616A-387
6	61	5.5	997	3	US-09-188-930-23
7	61	5.5	1816	3	US-09-188-930-262
8	48.8	4.4	1311	4	US-09-068-140A-9
9	48.8	4.4	5150	4	US-09-068-140A-14
10	47.6	4.3	7218	1	US-08-232-463-14
11	42.2	3.8	1594	2	US-08-933-750C-61
12	42.2	3.8	1594	3	US-09-234-613-61
13	40.8	3.7	1259	3	US-08-961-871-11
14	39.8	3.6	1192	4	US-09-142-565-1
15	39.6	3.6	897	4	US-09-434-288-6
16	39.2	3.5	1177	3	US-08-961-871-9
17	39	3.5	2634	1	US-08-196-218-31
18	39	3.5	2634	1	US-08-681-953-31
19	36.8	3.3	8460	1	US-08-469-005A-9
20	35.6	3.2	1535	1	US-08-910-973-10
21	35.6	3.2	1535	4	US-09-499-227-10
22	35.6	3.2	1550	3	US-09-234-332-3
23	35.6	3.2	1722	4	US-09-385-028-15
24	35.6	3.2	11604	4	US-09-385-028-13
25	35.6	3.2	15079	4	US-09-385-028-1
26	35.4	3.2	606	2	US-08-432-871C-48
27	35.4	3.2	606	4	US-09-270-956-48

28	35.4	3.2	2681	4	US-08-928-213B-7	Sequence 7, Appli
29	35.2	3.2	50341	1	US-08-247-901C-1	Sequence 1, Appli
30	35.2	3.2	50341	2	US-09-075-904-1	Sequence 1, Appli
31	35.2	3.2	52297	4	US-09-426-436-1	Sequence 1, Appli
32	35.2	3.2	52297	4	US-08-705-557-1	Sequence 1, Appli
33	35.2	3.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
34	34.8	3.1	1949	2	US-08-937-466-3	Sequence 3, Appli
35	34.8	3.1	1949	2	US-09-172-528-3	Sequence 3, Appli
36	34.8	3.1	1949	3	US-09-318-199-3	Sequence 3, Appli
37	34.8	3.1	1949	4	US-09-503-579-3	Sequence 3, Appli
38	34.6	3.1	1421	4	US-08-943-731-207	Sequence 207, App
39	34.6	3.1	20084	4	US-08-943-731-5	Sequence 5, Appli
40	34.2	3.1	34303	2	US-08-735-609-4	Sequence 4, Appli
41	34.2	3.1	34303	2	US-08-735-609-4	Sequence 4, Appli
42	34.2	3.1	34303	3	US-09-315-372-4	Sequence 4, Appli
43	34.2	3.1	34303	3	US-09-244-752-4	Sequence 4, Appli
44	34.2	3.1	34303	3	US-09-245-497-4	Sequence 4, Appli
45	34.2	3.1	34303	4	US-09-562-919-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-933-750C-68
; Sequence 68, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNCOAT01

CLONE: 724157
US 08-933-750C-68

Query Match 94.6%; Score 1054.4; DB 2; Length 1643;
Best Local Similarity 97.8%; Pred. No. 3.9e-264;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCTGCGGGATCAGCCCTCCAGCAAAATGGT 60
DB 192 CTGAAGCTTCAAGATGGCTGACAGGACCTGCGGGATCAGCCCTCCAGCAAAATGGT 251
QY 61 GGCCTCAGGACCGCGGCTGTGTTACTCTCTCTTCATGACACCCCTGGACGTGTGAA 120
DB 252 GGCCTCAGGACCGCGGCTGTGTTACTCTCTCTTCATGACACCCCTGGACGTGTGAA 311
QY 121 GGTTCGCTCGAGTCTCAGCGCCCTCCATGGCCAGCGAGTGTATGCTTCTCCAGACT 180
DB 312 GGTTCGCTCGAGTCTCAGCGCCCTCCATGGCCAGCGAGTGTATGCTTCTCCAGACT 371
QY 192 GTGAGGCTCTCTATACCAAAATGGCTCTCTCTTCCAAATCCAGGGAAGTCCCTCT 240
DB 372 GTGAGGCTCTCTATACCAAAATGGCTCTCTCTTCCAAATCCAGGGAAGTCCCTCT 407
QY 241 GTATTGAATGGTGTCTGGAGCTCTGTACTGTGCCAAATGGTGGCGCTGTGCCAC 300
DB 408 GTATTGAATGGTGTCTGGAGCTCTGTACTGTGCCAAATGGTGGCGCTGTGCCAC 467
QY 301 CTGTTTCAAGACCTTACCGCTTCACTGACCAATGATGCTTCTGTAAGATCTGTAG 360
DB 468 CTGTTTCAAGACCTTACCGCTTCACTGACCAATGATGCTTCTGTAAGATCTGTAG 527
QY 361 GCAGGAGGACAGGACCTCTGAGGCGCTCCCGCCAGCTGGGATGACTGTGCC 420
DB 528 GCAGGAGGACAGGACCTCTGAGGCGCTCCCGCCAGCTGGGATGACTGTGCC 587
QY 421 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTCGAGC 480
DB 588 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTCGAGC 647
QY 481 CTGACCTCTGACCTTACGACCCATGGTGGCTGCGGCTGGCCCGCTTGGGCAACCGT 540
DB 648 CTGACCTCTGACCTTACGACCCATGGTGGCTGCGGCTGGCCCGCTTGGGCAACCGT 707
QY 541 GACTGTGATCAGCCCTTGGAGCTTATGCGACAAGCTGAGGCTCAGCATGTGCTGA 600
DB 708 GACTGTGATCAGCCCTTGGAGCTTATGCGACAAGCTGAGGCTCAGCATGTGCTGA 767
QY 601 CCGGAGCTGGGTGCTGTGTGAACTGCAAGTGGCTCAGGCTGGCTGGCGCTCACTGTG 660
DB 768 CCGGAGCTGGGTGCTGTGTGAACTGCAAGTGGCTCAGGCTGGCTGGCGCTCACTGTG 827
QY 661 GCTGGGCTGGGCGCCATCTGCGCTTTCAGATGTGCGCTTCTCAGCCCTGTACTGTTCAA 720
DB 828 GCTGGGCTGGGCGCCATCTGCGCTTTCAGATGTGCGCTTCTCAGCCCTGTACTGTTCAA 887
QY 721 CTATGAGCTGGTGAAGAGCTGCTCAATGGCTCAGCGCGAGGACAGACTCTGTGG 780
DB 888 CTATGAGCTGGTGAAGAGCTGCTCAATGGCTCAGCGCGAGGACAGACTCTGTGG 947
QY 781 CATGAGCTTGTGGCTGTGGCTCTCAGGACGCTGGCTGCGAGTGTGACTTACCTT 840
DB 948 CATGAGCTTGTGGCTGTGGCTCTCAGGACGCTGGCTGCGAGTGTGACTTACCTT 1007
QY 841 TGACCTGGTAAAGACCCCAAGCTGGCTCTTGGAGCGATGGAGGTGTGAGGTGAA 900
DB 1008 TGACCTGGTAAAGACCCCAAGCTGGCTCTTGGAGCGATGGAGGTGTGAGGTGAA 1067
QY 901 CCCCCTGCATGTGAGTCACTGCTGCTGCTGCGAGGATCCGGGCGAGTGGGCGAC 960
DB 1068 CCCCCTGCATGTGAGTCACTGCTGCTGCTGCGAGGATCCGGGCGAGTGGGCGAC 1127
QY 961 CAAGGAGCTTCTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTCTGTGCAT 1020

DB 1128 CAAGGAGCTTCTTGGAGGCTTCTTCTTCGAGATCAAGGCTGGCCCTCTCTGTGCAT 1187
QY 1021 CATGATCAGCACCTATGACTTGGCAAAAGCTTCTCCAGAGGCTGAACCGAGCCGCT 1080
DB 1188 CATGATCAGCACCTATGACTTGGCAAAAGCTTCTCCAGAGGCTGAACCGAGCCGCT 1247
QY 1081 TCTGGGCGCTGAAGGCGCAAGGCAAGGCAAGGAC 1114
DB 1248 TCTGGGCGCTGAAGGCGCAAGGCAAGGCAAGGAC 1281

RESULT 2

US-09-234-613-68
; Sequence 68, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYN00AT01
; CLONE: 724157
; US-09-234-613-68

Query Match 94.6%; Score 1054.4; DB 3; Length 1643;
Best Local Similarity 97.8%; Pred. No. 3.9e-264;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCTTGGGGATCAGCCCTCCAGCAAAATGGT 60
DB 192 CTGAAGCTTCAAGATGGCTGACAGGACCTTGGGGATCAGCCCTCCAGCAAAATGGT 251
QY 61 GGCCTCAGGACCGCGGCTGTGTTACTCTCTTTCATGACACCCCTGGACGTGTGAA 120

Db 252 GGCCTCAGGACCGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGGAGGTGGTAA 311
 Qy 121 GGTTCGCTCGAGTCTCAGCGGCCCTCATGCGCAGCGAGTGTGCTCTCTCCAGACT 180
 Db 312 GGTTCGCTCGAGTCTCAGCGGCCCTCATGCGCAGCGAGTGTGCTCTCTCCAGACT 371
 Qy 181 GTGAGGCTCTCTATACAAATTTGGCCTCTCTCTCCAAATCCAGGAAAGTCTCTCT 240
 Db 372 GTGAGGCTCTCTATACAAATTTGGCCTCTCTCTCCAAATCCAGGAAAGTCTCTCT 407
 Qy 241 GTATTCAATGGTCTCTGAGGCTCTGTATCTGTGTCGCCAAATGTGCGCTGTGCCAC 300
 Db 408 GTATTCAATGGTCTCTGAGGCTCTGTATCTGTGTCGCCAAATGTGCGCTGTGCCAC 467
 Qy 301 CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGATGCTTCTGTGAAGATCTGTAG 360
 Db 468 CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGATGCTTCTGTGAAGATCTGTAG 527
 Qy 361 GCAGAGGACACAGGACCTCTGAGGCGCTCTGAGGCGCTCTCCCGCACCTGTGATGTGCG 420
 Db 528 GCAGAGGACACAGGACCTCTGAGGCGCTCTGAGGCGCTCTCCCGCACCTGTGATGTGCG 587
 Qy 421 AGCTACCGCATCTACTTCACTGCTATGACCACTGAGGCTTCTGTGTGTCGCTGAGC 480
 Db 588 AGCTACCGCATCTACTTCACTGCTATGACCACTGAGGCTTCTGTGTGTCGCTGAGC 647
 Qy 481 CTTGACTCTGACCTTACGACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 648 CTTGACTCTGACCTTACGACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707
 Qy 541 GACTGTGATCAGCCCTGAGCTTATGCGGCAAGAGTGTGAGJ.TCAGATGTGTGTA 600
 Db 708 GACTGTGATCAGCCCTGAGCTTATGCGGCAAGAGTGTGAGJ.TCAGATGTGTGTA 767
 Qy 601 CCGGAGCTGGGTGCTGTGTGTGAACTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTG 660
 Db 768 CCGGAGCTGGGTGCTGTGTGTGAACTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTG 827
 Qy 661 GTGAGCTGGGGGCCCCACTGCTTCTGAGATGTGCTTCTCAGCCCTGTACTGTGTTCAA 720
 Db 828 GTGAGCTGGGGGCCCCACTGCTTCTGAGATGTGCTTCTCAGCCCTGTACTGTGTTCAA 887
 Qy 721 CTATGAGCTGTGAGAGCTGTGCTTCAATGGCTCAGGCGGAGGACAGACTTCTGTGG 780
 Db 888 CTATGAGCTGTGAGAGCTGTGCTTCAATGGCTCAGGCGGAGGACAGACTTCTGTGG 947
 Qy 781 CATGAGCTTTGTGCTGTGTGCTTCTCAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 948 CATGAGCTTTGTGCTGTGTGCTTCTCAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1007
 Qy 841 TGACGTGTAAGACCAACGCGAGTGTGCTCTGAGGACGATGAGGCTGTGAGAGTGA 900
 Db 1008 TGACGTGTAAGACCAACGCGAGTGTGCTCTGAGGACGATGAGGCTGTGAGAGTGA 1067
 Qy 901 CCCCCTGATGTGAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 1068 CCCCCTGATGTGAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
 Qy 961 CAAGGACTCTTTGAGGCTTCTCTCTCGATCATCAAGGCTGCCCCCTCTCTGTCAT 1020
 Db 1128 CAAGGACTCTTTGAGGCTTCTCTCTCGATCATCAAGGCTGCCCCCTCTCTGTCAT 1187
 Qy 1021 CATGATCAGACCTATGAGTGTGGGAAAGTCTTCTCAGAGGCTGAACAGGACCGGCT 1080
 Db 1188 CATGATCAGACCTATGAGTGTGGGAAAGTCTTCTCAGAGGCTGAACAGGACCGGCT 1247
 Qy 1081 TCTGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 1114
 Db 1248 TCTGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 1281

RESULT 3

US-09-605-785-387
 ; Sequence 387, Application US/09605785
 ; Patent No. 6321716
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqi
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Hepler, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C16
 ; CURRENT APPLICATION NUMBER: US/09/605,785
 ; CURRENT FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 835
 ; SOFTWARE: FastSBQ for Windows Version 3.0
 ; SEQ ID NO 387
 ; LENGTH: 537
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-605-785-387

Query Match 15.3%; Score 170; DB 4; Length 537;
 Best Local Similarity 100.0%; Pred. No. 4.3e-35;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 945 GGGCCGAGTCGGGCACCAAGGAGCTCTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
 Db 1 GGGCCGAGTCGGGCACCAAGGAGCTCTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60

Qy 1005 CCCCCTCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAGCTTCTTCAGAGGC 1064
 Db 61 CCCCCTCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAGCTTCTTCAGAGGC 120

Qy 1065 TGACCCAGACCGGCTTCTGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
 Db 121 TGAACCCAGACCGGCTTCTGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 170

RESULT 4
 US-09-439-313-387
 ; Sequence 387, Application US/09439313
 ; Patent No. 6329505
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan Louise
 ; APPLICANT: Jiang Yuqi
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Kalos, Michael
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Retter, Mark
 ; APPLICANT: Soik, John
 ; APPLICANT: Day, Craig
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C9
 ; CURRENT APPLICATION NUMBER: US/09/439,313
 ; CURRENT FILING DATE: 1999-11-12
 ; NUMBER OF SEQ ID NOS: 575

[illegible]D_b 44

```

618 GTGTTGAAGCTGCAAGTGCTCAGGCTGGCTGGCGCTACTGTGCTGGGCTGGGGGCCCCA 677
Db 503 GCGTCAGGAGGGGTGTTGGTAGAAGAAGGATTTGGCGCAITTTTACGCATCTTATCGAACAA 562
Qy 678 CTGCCCTTCGAGATGTGCCCTTCTCAGCCCTCTACTGTTCAACTATGAGCTGGTGAAGA 737
Db 563 CTGTGTCATGATGCCCGGTTTACGGCGTTCACCTTCGCCACATATGAGCCACGAGA 622
Qy 738 GCTGGCTCAATGGGCTCAGGCCGAA-----GGACCAGACTTCTGTGGGCATGAGCTTTC 791
Db 623 AAGSGTTGTTGGAGGTGTCGCCGAGACTGCGAACGATGAGAAITTTGTTAGTCATGCTA 682
Qy 792 TGGCTGTGTCATCTCAGGAGCGTGGCTGAGTGCTGACTCTACCTTTCAGCTGGTAA 851
Db 683 CTGCTGTGTCGCTGCTGGAGCTTTGGCTGCGAGTAGTAACCACTCCACTAGATGTTGTCA 742
Qy 852 AGACCCMAAGCCAG 865
Db 743 AAACACTGATTGCAG 756

RESULT 9
US-09-068-140A-14
; Sequence 14, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,690
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:

```

REGISTRATION NUMBER: PCT/EP96/04807
 FILING DATE: NO. 6281409ember 4, 1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Dinner, Dara L.
 REGISTRATION NUMBER: 33,690
 REFERENCE/DOCKET NUMBER: C70237
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-5017
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
 LENGTH: 5150 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Ribes nigrum
 STRAIN: Ben Alder
 'S-09-068-140A-14

Query Match 4.4%; Score 48.8; DB 4; Length 5150;

Best Local Similarity 45.7%; Pred. No. 0.0024;
 Matches 253; Conservative 0; Mismatches 292; Indels 9; Gaps 2;

QY 318 CCGCTTCACTGCGCACCATGCGTCCCTGCGTGAAGATCGTGAGGACCGAGGCGACAGGA 377
 DB 3370 CTCGAATCCGCGGCTCCGACAGCCCTTGGTGCATACCTGAAGTTGAAGTCCCGCG 3429
 QY 378 CCCTCTGCGAGCGCCCTCCCGCCGCTGCGTGAAGTGTGCGAGCTACCGCCCTACT 437
 DB 3430 GACTTTACCGTGGCTTGGTGAAGTGTGCGTGAAGTGTGCGAGCTACCGCCCTACT 3489
 QY 438 TCAGTCTGATGACCAATGAAGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 497
 DB 3490 TCTCCGTTTACGAGATGTGTAAGGAGACTTTTCTCATGTGATCCGAGC---AATTCCG 3546
 QY 498 ACGCACCATGTTGGTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCT 557
 DB 3547 GTGCGCAGCGCTTTCGGGGGTGTTCCGACGCTGGCGGCTGGCGGCTGGCGGCTGGCGG 3606
 QY 558 TGGAGCTTATCGGCAAAAGCTGCAAGCTGATGTTGCTACCGGAGCTGGGTCCT 617
 DB 3607 TGGATGTGTGAACAGAGGTTGCGAGTTGCGAGCAGCTCGTACAAAGGCTGTTGTTGATT 3666
 QY 618 GTGTTGAACTGCACTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCT 677
 DB 3667 GCGTCAGGAGGCTGTTGGTAGAAGAGGATGGCGCATTTACGATCTTATCGAACAA 3726
 QY 678 CTGCGCTTCAGATGTCCTTCTCAGCCCTGTTACTGGTTCAACTATGAGTGGTGAAGA 737
 DB 3727 CTGTGTCATGATGCGCGCTTACGCGCTTACCTTCGCCCATATGAGCCAGGA 3786
 QY 738 GCTGGCTCAATGGGCTCAGCGCAA-----GACCAAGACTTCTGTGGGCATGAGCTTTG 791
 DB 3787 AAGGCTTGTGGAGGTGTCGCGGAGACTGCGCAACGATGAGAAATTTGTAGTGCATGTA 3846
 QY 792 TGGCTGTGGCATCTCAGGACGGTGGCTGAGTGTGCTGACTTACCTTTGACGTGGTAA 851
 DB 3847 CTGCTGTGCTGCTGCTGAGGCTTTGGCTGAGGCTTTGGCTGAGTGAACCTCACTAGATGTTGCA 3906

RESULT 10

US-08-232-463-14
 ; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: pTZgpt-Fls
 US-08-232-463-14

Query Match 4.3%; Score 47.6; DB 1; Length 7218;

Best Local Similarity 5.0%; Pred. No. 0.0055;

Matches 20; Conservative 213; Mismatches 167; Indels 0; Gaps 0;

QY 132 AGTCTACGGCCCTCCATGCGCAGGAGCTGATGCTTCTCCACACTGTGAGCCTCT 191
 DB 1030 ATTCGAGCTTGGCTCGAGCTCGAGGAGCTTGCATYYYYYYYYYYYYYYYYYY 1089
 QY 192 CCTATACAAATTCCTCTCTCTCCCAATCCACAGGAAAGTCCCTCTGATTCGAATG 251
 DB 1090 YY 1149
 QY 252 GTGCTCTGAGCCTCTGTACCTGTGCCCAATGTTGCCCTGTGCCACTGTTCAAG 311
 DB 1150 YY 1209
 QY 312 ACCCTACCGCTTCACTGSCACATGATGCTTCTGAAGATCGTAGGACAGGCA 371
 DB 1210 YY 1269
 QY 372 CCAGGACCTCTGAGCGGCTCCCGCCACCTGGTGTGATGACTGTCCAGTACCGCA 431
 DB 1270 YY 1329
 QY 432 TCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTTGTCGAGCCCTGACCTG 491
 DB 1330 YY 1389
 QY 492 ACCTCTACGACCCATGTTGGCTGGCGCTGGCCCGCT 531
 DB 1390 YY 1429

RESULT 11

US-08-933-750C-61

; Sequence 61, Application US/08933750C

; Patent No. 5932442

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

```

; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNNOT02
; CLONE: 207452
; US-08-933-750C-61

Query Match 3.8%; Score 42.2; DB 2; Length 1594;
Best Local Similarity 59.7%; Pred. No. 0.082;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 951 AGTCGGGCACCAAGGACTCTTTGCGAGGCTTCCTTCGCGATCATCAAGGCTGCCCT 1010
Db 1001 AGAAGGCGCCCTGGGCTCTTCAAGGGCTGTCGCCAGCTTGCTGAAGGCTGCCCTCT 1060

QY 1011 CTGTGCGCATCATGATCAGCACTATGATGTCGCAAAAGCTTCTCCAGAGGCTGAAC 1069
Db 1061 CCACAGGCTTCATGTTCTTCTCGATGAATTTCTGTATGTTCTTCCACTGCATGAAC 1119

RESULT 12
US-09-234-613-61
; Sequence 61, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.

```

```

; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNNOT02
; CLONE: 207452
; US-09-234-613-61

Query Match 3.8%; Score 42.2; DB 3; Length 1594;
Best Local Similarity 59.7%; Pred. No. 0.082;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 951 AGTCGGGCACCAAGGACTCTTTGCGAGGCTTCCTTCGCGATCATCAAGGCTGCCCT 1010
Db 1001 AGAAGGCGCCCTGGGCTCTTCAAGGGCTGTCGCCAGCTTGCTGAAGGCTGCCCTCT 1060

QY 1011 CTGTGCGCATCATGATCAGCACTATGATGTCGCAAAAGCTTCTCCAGAGGCTGAAC 1069
Db 1061 CCACAGGCTTCATGTTCTTCTCGATGAATTTCTGTATGTTCTTCCACTGCATGAAC 1119

RESULT 13
US-08-961-871-11
; Sequence 11, Application US/08961871
; Patent No. 6013858
; GENERAL INFORMATION:
; APPLICANT: Wallace, Douglas C.
; APPLICANT: Graham, Brett H.
; APPLICANT: MacGregor, Grant R.
; TITLE OF INVENTION: Mouse Lacking Heart-Muscle Adenine
; Nucleotide Translocator Protein and Methods
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winmer and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

Db 415 GCGGTCTCGAGGACACCCGCTCAGCGCGTCTCTCCACCTGCCGCCCAACCGTGACTCC 474
Qy 491 GACCTCTAGCAACCCATGGTGGCGCGCTGGCCCGCTGGGCAACCGTGACTGTGATC 550
Db 475 GAGCGGCTCGCGCGCACCGACCCGGAGGCACTGCCCGCGGTCTGTAAACCGGAAGGCCACC 534
Qy 551 AGCCCCCTGGAGCT 564
Db 535 GCCGCGCTGCACCT 548

```

Search completed: January 22, 2003, 19:05:41
 Job time : 75.6489 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 17:33:27 ; Search time 55.7 Seconds
(without alignments)
8985,296 million cell under

```
Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagattcaaatggtc.....agggcagaagccgaagcac 1114
```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Tr number of hits satisfying chosen parameters: 793544

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : Published_Applications_NA.**
1: /cgn2_6/pdataa1/pubpna/US07_PUBCOMB.seq:**
2: /cgn2_6/pdataa1/pubpna/PCT_NEW_PUB.seq:**
3: /cgn2_6/pdataa1/pubpna/US06_NEW_PUB.seq:**
4: /cgn2_6/pdataa1/pubpna/US06_PUBCOMB.seq:**
5: /cgn2_6/pdataa1/pubpna/US07_NEW_PUB.seq:**
6: /cgn2_6/pdataa1/pubpna/PCTUS_PUBCOMB.seq:**
7: /cgn2_6/pdataa1/pubpna/US08_NEW_PUB.seq:**
8: /cgn2_6/pdataa1/pubpna/US08_PUBCOMB.seq:**
9: /cgn2_6/pdataa1/pubpna/US09_NEW_PUB.seq:**
10: /cgn2_6/pdataa1/pubpna/US09_PUBCOMB.seq:**
11: /cgn2_6/pdataa1/pubpna/US10_NEW_PUB.seq:**
12: /cgn2_6/pdataa1/pubpna/US10_PUBCOMB.seq:**
13: /cgn2_6/pdataa1/pubpna/US06_NEW_PUB.seq:**
14: /cgn2_6/pdataa1/pubpna/US06_PUBCOMB.seq:**
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1114	100.0		1114	10	US-09-888-358-1	Sequence 1, Appli
2	1054.4	94.6		1546	10	US-09-888-358-2	Sequence 2, Appli
3	1054.4	94.6		1843	10	US-09-840-787-68	Sequence 68, Appli
4	185.8	16.7		453	10	US-09-867-701-4486	Sequence 4486, Ap
c	170.2	15.3		512	10	US-09-917-800A-329	Sequence 329, App
6	170	15.3		537	9	US-10-012-896-387	Sequence 387, App
7	170	15.3		537	9	US-09-895-793-387	Sequence 387, App
8	170	15.3		537	9	US-09-895-814-387	Sequence 387, App
9	170	15.3		537	10	US-09-759-143-387	Sequence 387, App
10	170	15.3		537	10	US-09-780-669-387	Sequence 387, App
11	170	15.3		537	10	US-09-822-827-387	Sequence 387, App
12	154	13.8		275	10	US-09-908-711-46	Sequence 46, Appli
13	116.2	10.4		246	10	US-09-960-352-2543	Sequence 2543, Ap
c	93.4	8.4		452	10	US-09-960-107-1133	Sequence 1133, Ap
c	74.8	6.7		452	10	US-09-864-761-11674	Sequence 11674, A
15	74.8	6.7		496	10	US-09-864-761-11708	Sequence 11708, A
16	60	5.4		3334	9	US-09-992-598-288	Sequence 288, App
17	60	5.4		3334	9	US-09-989-293A-288	Sequence 288, App
18	60	5.4		3334	9	US-10-063-547-57	Sequence 57, Appli
19	60	5.4		3334	9		

ALIGNMENTS

```

RESULT 1
US-09-888-358-1
; Sequence 1, Application US/09888358
; Patent No. US20020119137A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID
; APPLICANT: ADAMS, SEAN H.
; APPLICANT: XU, XING XIAN
; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 10716/66
; CURRENT APPLICATION NUMBER: US/09/888,358
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,307
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-358-1

```

Query Match	100.0%;	Score 1114;	DB 10;	Length 1114;
Best Local Similarity	100.0%;	Pred. No. 7e-286;		
Matches 1114;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps
1 CTGAGCCTTCAAGATGCTGACACAGACCCCTGGCGCATCAGCCCTCCAGCAAAATGCT	60			
1 CTGAGCCTTCAAGATGCTGACACAGACCCCTGGCGCATCAGCCCTCCAGCAAAATGCT	60			
61 GGCTTAGGCAACGGGCGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGACCTGGTGAA	120			
61 GGCTTAGGCAACGGGCGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGACCTGGTGAA	120			
121 GGTTCGCTGCAGTCTCAGGGCCCTCATGGCCAGGAGCTGATGCTTCTCCAGACT	180			
121 GGTTCGCTGCAGTCTCAGGGCCCTCATGGCCAGGAGCTGATGCTTCTCCAGACT	180			
181 GTGGAGCCTCTCTTATACAAATTCGCCCTCTCTCTTCCAAATCCACAGGGAAGTGCCTCT	240			
181 GTGGAGCCTCTCTTATACAAATTCGCCCTCTCTCTTCCAAATCCACAGGGAAGTGCCTCT	240			

QY 241 GTATTGAATGGTCTCTGAGGCTCTGTACTGTGCCAAATGGTGGCCGCTGGCCAC 300
 Db 241 GTATTGAATGGTCTCTGAGGCTCTGTACTGTGCCAAATGGTGGCCGCTGGCCAC 300
 QY 301 CTGGTTTCAAGACCCCTACCGCTTCACTGGCAACATGATGCTTCTGTAAGATCGTGAG 360
 Db 301 CTGGTTTCAAGACCCCTACCGCTTCACTGGCAACATGATGCTTCTGTAAGATCGTGAG 360
 QY 361 GCACGAGGACACAGGACCCCTCTGGAGGGCTTCCCGCCACCCCTGGTGTGATGTGCC 420
 Db 361 GCACGAGGACACAGGACCCCTCTGGAGGGCTTCCCGCCACCCCTGGTGTGATGTGCC 420
 QY 421 AGCTACGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTTGTGTGAGC 480
 Db 421 AGCTACGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTTGTGTGAGC 480
 QY 481 CTTGACCTCTGACCTCTACCGACCCATGGTGGCTGGGGGCTGGCCGCTGGGACCGT 540
 Db 481 CTTGACCTCTGACCTCTACCGACCCATGGTGGCTGGGGGCTGGCCGCTGGGACCGT 540
 QY 541 GACTGTGATCAGCCCTCTGAGGCTTATCGGACAAAGCTGCAAGCTCAGAGCTGAGTGTGTA 600
 Db 541 GACTGTGATCAGCCCTCTGAGGCTTATCGGACAAAGCTGCAAGCTCAGAGCTGAGTGTGTA 600
 QY 601 CCGGAGCTGGTGGCTGTGTTGCACTGCACTGAGTGGCTCAGGCTGGCTGGCTCACTGTG 660
 Db 601 CCGGAGCTGGTGGCTGTGTTGCACTGCACTGAGTGGCTCAGGCTGGCTGGCTCACTGTG 660
 QY 661 GCTGGGCTGGGGCCCACTGCTTCCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
 Db 661 GCTGGGCTGGGGCCCACTGCTTCCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
 QY 721 CTATGAGCTGGTGAAGCTGGCTCAATGGCTCAGGCCGAAGGACCACTTCTGTGGG 780
 Db 721 CTATGAGCTGGTGAAGCTGGCTCAATGGCTCAGGCCGAAGGACCACTTCTGTGGG 780
 QY 781 CATGAGCTTTTGGCTGGTGGCTCTCAGGACCGTGGCTGCACTGCTGACTTACCTT 840
 Db 781 CATGAGCTTTTGGCTGGTGGCTCTCAGGACCGTGGCTGCACTGCTGACTTACCTT 840
 QY 841 TGAGTGGTAAAGACCAACGCGCTGCTCTGGAGCGCATGGAGGCTGTGAGAGTGA 900
 Db 841 TGAGTGGTAAAGACCAACGCGCTGCTCTGGAGCGCATGGAGGCTGTGAGAGTGA 900
 QY 901 CCCCTGCATGTGGAATCCACTGCTGCTGCGGAGATCCGGGCGGAGTGGGAC 960
 Db 901 CCCCTGCATGTGGAATCCACTGCTGCTGCGGAGATCCGGGCGGAGTGGGAC 960
 QY 961 CAAGGACTCTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTTGCCAT 1020
 Db 961 CAAGGACTCTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTTGCCAT 1020
 QY 1021 CATGATCAGCACCTATGAGTTCGGCAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1080
 Db 1021 CATGATCAGCACCTATGAGTTCGGCAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1080
 QY 1081 TCTGGGCGCTGAAGAGGCGAAGAGGCAAGGAC 1114
 Db 1081 TCTGGGCGCTGAAGAGGCGAAGAGGCAAGGAC 1114

RESULT 2

US-09-888-358-2
 ; Sequence 2, Application US/09888358
 ; Patent No. US202019137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LEWIN, DAVID
 ; APPLICANT: ADAMS, SEAN H.
 ; APPLICANT: YU, XING XIAN
 ; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
 ; FILE REFERENCE: 10716/56
 ; CURRENT APPLICATION NUMBER: US/09/888,358
 ; CURRENT FILING DATE: 2001-06-22

; PRIOR APPLICATION NUMBER: 60/213,307
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1546
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-888-358-2
 Query Match 94.6%; Score 1054.4; DB 10; Length 1546;
 Best Local Similarity 97.8%; Pred. No. 4.9e-270;
 Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
 QY 1 CTGAAGCTTCAAGATGGCTACACGAGACCTTGGGGGCTATCAGCCGCTCCAGCAAAATGGT 60
 Db 105 CTGAAGCTTCAAGATGGCTACACGAGACCTTGGGGGCTATCAGCCGCTCCAGCAAAATGGT 164
 QY 61 GGCCTCAGGACACCGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGAGCTGGTGA 120
 Db 165 GGCCTCAGGACACCGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGAGCTGGTGA 224
 QY 121 GGTTCGCTCGAGTCTCAGCGCCCTCCATGGCCAGCGAGCTGATGCCCTTCTCCAGACT 180
 Db 225 GGTTCGCTCGAGTCTCAGCGCCCTCCATGGCCAGCGAGCTGATGCCCTTCTCCAGACT 284
 QY 181 GTGAGGCTCTCTTATACCAATTCGCCCTCTCTTCCAAATCCACAGGAGAGTGCCTCT 240
 Db 285 GTGAGGCTCTCTTATACCAATTCGCCCTCTCTTCCAAATTCGCCCTCTCTTCCAAAT 320
 QY 241 GTATGCAATGGTGTCTCTGGAGCTCTGTACTGTGCCAAATGGTGGCCGCTGTGCCAC 300
 Db 321 GTATGCAATGGTGTCTCTGGAGCTCTGTACTGTGCCAAATGGTGGCCGCTGTGCCAC 380
 QY 301 CTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATGGATGCTTCTGTAAGATCGTGAG 360
 Db 381 CTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATGGATGCTTCTGTAAGATCGTGAG 440
 QY 361 GCACGAGGACACAGGACCCCTCTGAGGCGGCTTCCCGCCACCCCTGGTGTGATGTGCC 420
 Db 441 GCACGAGGACACAGGACCCCTCTGAGGCGGCTTCCCGCCACCCCTGGTGTGATGTGCC 500
 QY 421 AGCTACGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTTGTGTGAGC 480
 Db 501 AGCTACGCCATCTACTTCACTGCTTATGACCACTGAAGGCTTCTGTTGTGTGAGC 560
 QY 481 CTTGACCTCTGACTCTACCGCACCATGGTGGCTGGGGGCTGGCCGCTGGGACCGT 540
 Db 561 CTTGACCTCTGACTCTACCGCACCATGGTGGCTGGGGGCTGGCCGCTGGGACCGT 620
 QY 541 GACTGTGATCAGCCCTCTGGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTGTA 600
 Db 621 GACTGTGATCAGCCCTCTGGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTGTA 680
 QY 601 CCGGAGCTGGTGGCTGTGTTGCACTGCACTGAGTGGCTCAGGCTGGCTGGCTCACTGTG 660
 Db 681 CCGGAGCTGGTGGCTGTGTTGCACTGCACTGAGTGGCTCAGGCTGGCTGGCTCACTGTG 740
 QY 661 GGTGGCTGGGGCCCACTGCTTCCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAA 720
 Db 741 GGTGGCTGGGGCCCACTGCTTCCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAA 800
 QY 721 CTATGAGCTGGTGAAGTGTGCTCAATGGGCTCAGGCCGAAGGACCACTTCTGTGGG 780
 Db 801 CTATGAGCTGGTGAAGTGTGCTCAATGGGCTCAGGCCGAAGGACCACTTCTGTGGG 860
 QY 781 CATGAGCTTTTGGCTGGTGGCTCTCAGGAGCGTGGCTGCACTGCTCAGCTTACCCTT 840
 Db 861 CATGAGCTTTTGGCTGGTGGCTCTCAGGAGCGTGGCTGCACTGCTCAGCTTACCCTT 920
 QY 841 TGACCTGTAAAGACCAACGCGCTGCTCTGGAGCGATGGAGGCTGTGAGAGTGA 900
 Db 921 TGACCTGTAAAGACCAACGCGCTGCTCTGGAGCGATGGAGGCTGTGAGAGTGA 980

Qy	901	CCCCCTGCATGTGGACTCCACCTTGGCTGGCTCTGCGAGGATCGGGCCGAGTCGGGCAC	960
Db	981	CCCCCTGCATGTGGACTCCACCTTGGCTGGCTCTGCGAGGATCGGGCCGAGTCGGGCAC	1040
Qy	961	CAAGGAGCTCTTTTCAGGCTTCCTTCTCGGATCATCAAGGCTGCCCTCCTGTGGCAT	1020
Db	1041	CAAGGAGCTCTTTTCAGGCTTCCTTCTCGGATCATCAAGGCTGCCCTCCTGTGGCAT	1100
Qy	1021	CATGATCAGGACCTATGAGTTTCGGCAAAGCTTCTTCCAGAGCTCAACACGAGCCGGCT	1080
Db	1101	CATGATCAGGACCTATGAGTTTCGGCAAAGCTTCTTCCAGAGCTCAACACGAGCCGGCT	1160
Qy	1081	TCTGGCGGCTTGAAGGGCAAGGGCAAGGAC	1114
Db	1161	TCTGGCGGCTTGAAGGGCAAGGGCAAGGAC	1194

RESULT 3

US-09-840-787-68
;
; ence 68, Application US/09840787
; ent No. US20020058264A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Bandman, Olga
; Shah, Purvi
; Au-Young, Janice
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSED: Incyte Pharmaceuticals, Inc.

Query Match 94.6%; Score 1054.4; DB 10; Length 1643;

Db 1188 CATGATCAGCACTATGAGTTCGGCAAAAGCTTCTCCAGAGCTGAACAGGACCGGCT 1247

QY 1081 TCTGGCGGCTGAAAGGGGCAAGGCAAGGAC 1114
Db 1248 TCTGGCGGCTGAAAGGGGCAAGGCAAGGAC 1281

RESULT 4

US-09-867-701-4486
; Sequence 4486, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Agiate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4486
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-4486

Query Match 16.7%; Score 185.8; DB 10; Length 453;
Best Local Similarity 92.0%; Pred. No. 4.3e-40;
Matches 207; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 503 CCCATGTTGGTGGCGGCTGGCGGCTGGCGCACCGTGTGATCAGCCCCCTGGAG 562
Db 230 CCCATTGTACCCCAATCCACACAGTGGGACCGTGTGATCAGCCCCCTGGAG 289
QY 563 CTTATGGCGCAAAAGCTGACGGCTAGCATGTGTGACCGGAGCTGGGTGCCTGTGTT 622
Db 290 CTTATGGCGCAAAAGCTGACGGCTAGCATGTGTGACCGGAGCTGGGTGCCTGTGTT 349
QY 623 CGAAGCTGAGGCTCAGGTGGCTGGCGCTCAGTGTGGCTGGGCGCCCACTGCC 682
Db 350 CGAAGCTGAGGCTCAGGTGGCTGGCGCTCAGTGTGGCTGGGCGCCCACTGCC 408
QY 683 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 727
Db 409 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 453

RESULT 5

US-09-917-800A-329/c
; Sequence 329, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336

; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 329
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA926365
; NAME/KEY: misc_feature
; LOCATION: (1)..(512)
; OTHER INFORMATION: n = a o r c o r t
US-09-917-800A-329

Query Match 15.3%; Score 170.2; DB 10; Length 512;
Best Local Similarity 81.7%; Pred. No. 6.1e-36;
Matches 196; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 857 CAACGCCAGGTGCTCTGGGAGCGATGGAGGCTGTGAGAGTGAACCCCTGCATGTGAC 916
Db 511 CAGCACAAATGTCACCTGGGAGCAGTGGAGGCTATGAGATGAGCGCCCAAGATTGAC 452
QY 917 TCACCTGGCTGTGCTGGGAGGATCGGGCCGAGTGGGCGCACCAAGGAGCTCTTTGCA 976
Db 451 TCACCTGGCTGTGCTGGGAGATCCAGGCTGAATCTGGCACCAAGGAGCTCTTTGCA 392
QY 977 GCCTTCCTCTCTCGGATCATCAAGGCTGCCCTCTCTGTCATCATGATCAGCACTAT 1036
Db 391 GGTTCCTCTCCNCGAGATCAAGGCTGGCGCTCTCTGCGCATCATGATCAGCACTAT 332
QY 1037 GAGTTGGCAAAAGCTT TCCAGAGGCTGAACAGGACCGGCTTCTGGCGGCTGAAAG 1096
Db 331 GAGTTGGCAAAAGCTTCTTCCACAGGCTCAACAGGAGCAGGCTCTGGGCCACTGAAGG 272

RESULT 6

US-10-012-896-387
; Sequence 387, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Fasih A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Mesgher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-387

Query Match 15.3%; Score 170; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 7,1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 120
QY 1065 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
Db 121 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 170

RESULT 7

US-09-895-793-387
; Sequence 387, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Basols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Basols, Carlota

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-387

Query Match 15.3%; Score 170; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 7,1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 120
QY 1065 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
Db 121 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 170

RESULT 8

US-09-895-814-387
; Sequence 387, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Basols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-387

Query Match 15.3%; Score 170; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 7,1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 120
QY 1065 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
Db 121 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 170

RESULT 9

US-09-759-143-387
; Sequence 387, Application US/09759143


```

; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-387

Query Match      15.3%; Score 170; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 7.1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60

QY 1005 CCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 120

QY 1065 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGCAAGGAGGCAAGGAC 1114
Db 121 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGCAAGGAGGCAAGGAC 170

RESULT 10
US-09-780-669-387
; Sequence 387, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.

```

```

; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-387

Query Match      15.3%; Score 170; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 7.1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60

QY 1005 CCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 120

QY 1065 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGCAAGGAGGCAAGGAC 1114
Db 121 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGCAAGGAGGCAAGGAC 170

RESULT 11
US-09-822-827-387
; Sequence 387, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-387

Query Match      15.3%; Score 170; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 7.1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60

QY 1005 CCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 120

QY 1065 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGCAAGGAGGCAAGGAC 1114
Db 121 TGAACCAAGACCGGCTTCTGGCGGCTGAAAGGGCAAGGAGGCAAGGAC 170

RESULT 12
US-09-908-711-46
; Sequence 46, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

```

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PA128
;; CURRENT APPLICATION NUMBER: US/09/908,711
;; CURRENT FILING DATE: 2001-07-20
;; PRIOR APPLICATION NUMBER: US01/01360
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,867
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01344
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,892
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01345
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,888
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01329
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,905
;; PRIOR FILING DATE: 2001-01-17
;; OR APPLICATION NUMBER: US01/01354
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,891
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01339
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,869
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01340
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,874
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01334
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,898
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01320
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,853
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01349
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,902
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01239
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,870
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01348
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,882
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01347
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,896
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01307
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,864
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01341
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,856
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01336
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,868
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01312
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 60/179,065
;; PRIOR FILING DATE: 2000-01-31
;; PRIOR APPLICATION NUMBER: 60/180,628

;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: 60/209,467
;; PRIOR FILING DATE: 2000-06-07
;; NUMBER OF SEQ ID NOS: 167
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 46
;; LENGTH: 275
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (246)
;; OTHER INFORMATION: n equals a,t,g, or c
US-09-908-711-46

Query Match 13.8%; Score 154; DB 10; Length 275;
Best Local Similarity 89.7%; Pred. No. 9.6e-32;

Matches 174; Conservative 2; Mismatches 17; Indels 1; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACCAGGACCCCTGCGGGCATCAGCCCTCCAGCAATGGT 60
|||||
Db 53 CTGAAGCTTCAAGATGGCTGACCAGGACCCCTGCGGGCATCAGCCCTCCAGCAATGGT 111
QY 61 GGCCTCAGGACACCGGGCTGTGGTTACCTCTCTTTCATGACACCCCTCGAGTGGTGAA 120
|||||
Db 112 GGCCTCAGGACACCGGGCTGTGGTTACCTCTCTTTCATGACACCCCTCGAGTGGTGAA 171
QY 121 GGTTCGCTCGAGTCTCAGCGGCCCTTCATGCGCAGGAGCTGATGCTTCTCCAGACT 180
|||||
Db 172 AGTTGCTCGAGTCTCAGCGGCCCTTCATGCGCAGGAGCTGATGCTTCTCCAGACT 231
QY 181 GTGGAGCCTCTCCT 194
|||||
Db 232 CTGGAGCCTCTCCT 245

RESULT 13

US-09-960-352-2543
; Sequence 2543, Application US/09960352
; Patent No. US20020137139A1

;; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan

;; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
;; FILE REFERENCE: 16511.006/37-21(10298)C
;; CURRENT APPLICATION NUMBER: US/09/960,352
;; CURRENT FILING DATE: 2001-09-24
;; NUMBER OF SEQ ID NOS: 15112
;; SEQ ID NO 2543
;; LENGTH: 246
;; TYPE: DNA
;; ORGANISM: Bos taurus

;; OTHER INFORMATION: Clone ID: 11-LIB34-070-Q1-E1-C11
US-09-960-352-2543

Query Match 10.4%; Score 116.2; DB 10; Length 246;
Best Local Similarity 87.9%; Pred. No. 9.6e-22;

Matches 138; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACCAGGACCCCTGCGGGCATCAGCCCTCCAGCAATGGT 60
|||||
Db 55 CTGAAGCTTCAAGATGGCTGACCAGGACCCCTGCGGGCATCAGCCCTCCAGCAATGGT 124
QY 61 GGCCTCAGGACACCGGGCTGTGGTTACCTCTCTTTCATGACACCCCTCGAGTGGTGAA 120
|||||
Db 125 GGCATCGCGCGGGGCTGTGGTCACTCCCTCTTCATGACCCCTCGATGGTGAA 184
QY 121 GGTTCGCTCGAG-TCTCAGCGGCCCTTCATGGCCAG 156
|||||
Db 185 GGTACGCTTGCAGTTCTCAGCGGCCCTCAGTGGCCAG 221

```
RESULT 14
US-09-880-107-1133/c
; Sequence 1133, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1133
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA450247
US-09-880-107-1133

Query Match      8.4%; Score 93.4; DB 10; Length 452;
Best Local Similarity 98.9%; Pred. No. 1.3e-15;
Matches 94; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1020 TCATGATCAGCACCTATCAGTTCGCGAAAGCTTCTTCAGAGGCTGAACAGGACCGGC 1079
Db 452 TCATGATCAGCACTATCAGTTCGCGAAGCTTCTTCAGAGGCTGAACAGGACCGGC 393

Qy 1080 TTTCTGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 1114
Db 392 TTTCTGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 358

RESULT 15
US-09-864-761-11674/c
; Sequence 11674, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11674
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004958.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
US-09-864-761-11674

Query Match      6.7%; Score 74.8; DB 10; Length 452;
Best Local Similarity 64.4%; Pred. No. 1.1e-10;
Matches 112; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 531 TGGGCACCGTGACTGTGATCAGCCCTCGAGCTTATGGGACAAAGCTCAGGCTCAGC 590
Db 436 TTGGTCAGTAAGTGTGATAGTCCACTAGATTGATTAGAACCAAGATGAGTCCAAGA 377

Qy 591 ATGTGTCGTACCGGAGCTGGGTGCCTGTGTTCGAACCTGCAGTGCCTCAGGCTGGCTGC 650
Db 376 AGTTTCTTACGTGGAAGTGCATCGATTGTTCAGCAAGAAAGTATCTGAAGATGGTTGA 317

Qy 651 GCTCACTGTGCTGGGCTGGGCGGCCCACTGCCCTTCGAGATGTCCTTCTTCAG 704
Db 316 TTTCCCTTTTGAGGGGCTGGGCTCCTACTGTCTTTAGAGATGTACTTCTTCAG 263
```

Search completed: January 22, 2003, 19:07:54
Job time : 70.7 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 19:06:17 ; Search time 1752.67 Seconds
(without alignments)
10293.906 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 cgaagcttcaagatgctg.....aggcgcaaggaggcaaggac 1114

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Tc number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estab.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estcov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1002.4	90.0	1402	11 AF119864	AF119864 Homo sapi
2	919.6	82.5	1020	13 BM554056	BM554056 AGENCOURT
3	848.6	76.2	1450	11 AK007934	AK007934 Mus muscu
4	831.8	74.7	1371	11 AK007276	AK007276 Mus muscu
5	826	74.1	891	9 AL522522	AL522522 AL522522
6	826	74.1	900	14 BQ883556	BQ883556 AGENCOURT

7	817.6	73.4	917	9	AL544360	AL544360
8	812.6	72.9	1453	11	AK002381	AK002381 Mus muscu
9	803.8	72.2	920	14	BQ683596	BQ683596 AGENCOURT
10	801	71.9	1491	11	AK019396	AK019396 Mus muscu
11	798.8	71.7	981	14	BM923995	BM923995 AGENCOURT
12	787	70.6	1001	9	AL519987	AL519987 AL519987
13	786.4	70.6	1029	13	BM459839	BM459839 AGENCOURT
14	785.6	70.5	921	9	AL529906	AL529906 AL529906
15	780.6	70.1	1012	13	BM468665	BM468665 AGENCOURT
16	777.8	69.8	1071	13	BI768227	BI768227 603056685
17	777.2	69.8	969	13	BI759919	BI759919 603044343
18	776.6	69.7	925	9	AL516722	AL516722 AL516722
19	775.4	69.6	911	9	AL528266	AL528266 AL528266
20	774.2	69.5	1061	14	BM923251	BM923251 AGENCOURT
21	766.4	68.8	816	12	BQ470240	BQ470240 602533758
22	764	68.6	885	9	AL514846	AL514846 AL514846
23	759.6	68.2	959	14	BM915493	BM915493 AGENCOURT
24	755.6	67.8	939	14	BQ949458	BQ949458 AGENCOURT
25	755.4	67.8	965	14	BQ934841	BQ934841 AGENCOURT
26	753.8	67.7	862	12	BQ470248	BQ470248 602533766
27	748.8	67.2	962	14	BQ688847	BQ688847 AGENCOURT
28	747.4	67.1	963	14	BQ071835	BQ071835 AGENCOURT
29	746.4	67.0	887	14	BQ954108	BQ954108 AGENCOURT
30	744.8	66.9	1474	11	AK010272	AK010272 Mus muscu
31	744.2	66.8	1077	14	BM808978	BM808978 AGENCOURT
32	743.2	66.7	875	9	AL555763	AL555763 AL555763
33	742	66.6	1043	14	BQ073154	BQ073154 AGENCOURT
34	738.6	66.3	847	9	AL547590	AL547590 AL547590
35	737	66.2	792	12	BQ470228	BQ470228 602533742
36	737	66.2	901	9	AL545150	AL545150 AL545150
37	735	66.0	976	14	BQ647782	BQ647782 AGENCOURT
38	734.2	65.9	961	14	BQ647042	BQ647042 AGENCOURT
39	733.4	65.8	1120	13	BM464148	BM464148 AGENCOURT
40	731.4	65.7	1080	14	BM917004	BM917004 AGENCOURT
41	729.2	65.5	918	13	BM457688	BM457688 AGENCOURT
42	727.8	65.3	890	9	AL515130	AL515130 AL515130
43	727.2	65.3	943	13	BM018915	BM018915 603646853
44	724.8	65.1	810	12	BQ749617	BQ749617 602707589
45	722.4	64.8	994	13	BM555499	BM555499 AGENCOURT

ALIGNMENTS

RESULT 1
AF119864
LOCUS AF119864 Homo sapiens PRO2163 mRNA, complete cds.
DEFINITION AF119864
ACCESSION AF119864
VERSION AF119864.1 GI:7770164
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhang, Y., Zhou, G., Bi, J., Liu, M. and He, F.
TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1402)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhang, Y., Zhou, G., Bi, J., Liu, M. and He, F.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
FEATURES
source
1. .1402
/organism="Homo sapiens"
/db_xref="taxon:9606"
/c1one="FLB8041"

```

/tissue_type="liver"
/dev_stage="fetus"
4..1017
/notes=predicted protein of H02163
/codon_start=1
/product="PRO2163"
/protein_id="AAF69618.1"
/db_xref="GI:7770165"
/translation="MVASGTVGVVTSLEMTPLDVVVKVLPQSRPMSSELSRLWS
LSYTWKCLLVGVNGLPPLVLCNGARCAWFQDPTFTGMDAFVKI VRHGGTTLW
ELMRPLVNTVPAITVFTAYDQKAFGLCGRAITSDVYAPWAGALRGITVWISPL
SGMRTKLQAHVSREYLGACVRAVAGQWRSILNIGGPTALRDVPSALYVVELV
KSNMLGRPKQTSVGVMSFVAGISRTVAALVLPFDVVKTORVALGAEAVRVNPL
HVDSTWLLLRIRAESQTKGLFAGFLPRLIKAAPSCAINTVFEFGKSFQRLNDEL
LGG"
BASE COUNT 261 a 432 c 401 g 308 t
ORIGIN

Query Match 90.0%; Score 1002.4; DB 11; Length 1402;
Best Local Similarity 97.6%; Pred. No. 9.2e-233;
Matches 1037; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 53 CAATGTGTGCTCAGGCAACGGGCTGTGTGTACCTCTCTTCTATGACACCCCTGGAC 112
DB 1 CAATGTGTGCTCAGGCAACGGGCTGTGTGTACCTCTCTTCTATGACACCCCTGGAC 60
QY 113 GTGTGAAGTTTGGCTGTCAGTCTCAGCGGCTCCATGGCCAGGAGCTGATCCTTCC 172
DB 61 GTGTGAAGTTTGGCTGTCAGTCTCAGCGGCTCCATGGCCAGGAGCTGATCCTTCC 120
QY 173 TCCAGACTGTGAGGCTCTCTCTATACCAAAATGGCCCTCTCTCTCAATCCACAGGGAAG 232
DB 121 TCCAGACTGTGAGGCTCTCTCTATACCAAT-----GGAAG 156
QY 233 TGCTCTGTATGAATGGTGTCTGAGGCTCTGTATCTGTGCCCAAAATGGTCCCGC 292
DB 157 TGCTCTGTATGAATGGTGTCTGAGGCTCTGTATCTGTGCCCAAAATGGTCCCGC 216
QY 293 TGTGCCACCTGTCTTCAAGACCTTACCGGCTTCACTGGCACCACCTGATCTCTGGAG 352
DB 217 TGTGCCACCTGTCTTCAAGACCTTACCGGCTTCACTGGCACCACCTGATCTCTGGAG 276
QY 353 ATGTGAGGACGAGGACACGAGACCTCTGAGCGGCTCTCCCGCCACCTCTGTGTATG 412
DB 277 ATGTGAGGACGAGGACACGAGACCTCTGAGCGGCTCTCCCGCCACCTCTGTGTATG 336
QY 413 ACTGTGCGAGCTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGTGT 472
DB 337 ACTGTGCGAGCTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGTGT 396
QY 473 GGTGAGGCTTACCTCTGACCTTACGACACCATGGTGGCTGGCGGCTGGCCCGCTG 532
DB 397 GGTGAGGCTTACCTCTGACCTTACGACACCATGGTGGCTGGCGGCTGGCCCGCTG 456
QY 533 GGCACCGTGTGATGACGCCCCCTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCAT 592
DB 457 GGCACCGTGTGATGACGCCCCCTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCAT 516
QY 593 GTGTGTACCGGAGCTGGGTGCTGTGTGCAACTGAGTGGCTCAGGCTGCTGGGCG 652
DB 517 GTGTGTACCGGAGCTGGGTGCTGTGTGCAACTGAGTGGCTCAGGCTGCTGGGCG 576
QY 653 TCACCTGTGCTGGGTGGGCCCCCCTGACCTTCAGATGTGCTTCTCAGGCTGTGAC 712
DB 577 TCACCTGTGCTGGGTGGGCCCCCCTGACCTTCAGATGTGCTTCTCAGGCTGTGAC 636
QY 713 TGTTCACATGAGCTGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACAGACT 772
DB 637 TGTTCACATGAGCTGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACAGACT 696
QY 773 TCTGTGGGATGAGCTTTGTGGCTGTGGCTGTCAGGAGCGGTGGCTGAGTGTGACT 832
DB 697 TCTGTGGGATGAGCTTTGTGGCTGTGGCTGTCAGGAGCGGTGGCTGAGTGTGACT 756

```

```

QY 833 CTACCTTTGACGTGTAAGACCCAAAGCCAGGTCGCTCTGGAGCGATGAGGCTGTG 892
DB 757 CTACCTTTGACGTGTAAGACCCAAAGCCAGGTCGCTCTGGAGCGATGAGGCTGTG 816
QY 893 AGAGTGAACCCCTTGCATGTGAGCTCCACCTGGCTGTCTGGAGGATCGGGCCGAG 952
DB 817 AGAGTGAACCCCTTGCATGTGAGCTCCACCTGGCTGTCTGGAGGATCGGGCCGAG 876
QY 953 TCCGCGCACCAAGGAGCTCTTTCAGAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCTTCC 1012
DB 877 TCCGCGCACCAAGGAGCTCTTTCAGAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCTTCC 936
QY 1013 TGTGCTCATGATGATCAGCACTTATGAGTTCCGAAAAGCTTCTTCCAGAGGCTGACCGAG 1072
DB 937 TGTGCTCATGATGATCAGCACTTATGAGTTCCGAAAAGCTTCTTCCAGAGGCTGACCGAG 996
QY 1073 GACCGGCTTCTGGCGGCTGAAAAGGGGCAAGGAGGCAAGGAC 1114
DB 997 GACCGGCTTCTGGCGGCTGAAAAGGGGCAAGGAGGCAAGGAC 1038

```

```

RESULT 2
BM554056 1020 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT.6546709 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742317
5', mRNA sequence.
ACCESSION BM554056
VERSION BM554056.1 GI:18793328
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1020)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloned Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12760 Row: e Column: 06
High quality sequence stop: 691.
Location/Qualifiers
1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5742317"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH MGC Library."
BASE COUNT 175 a 327 c 302 g 213 t 3 others
ORIGIN

```

```

FEATURES
Source
1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5742317"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH MGC Library."
BASE COUNT 175 a 327 c 302 g 213 t 3 others
ORIGIN
Query Match 82.5%; Score 919.6; DB 13; Length 1020;
Best Local Similarity 98.5%; Pred. No. 1e-212;
Matches 958; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

```

QY 1 CTGAAGCTTCAAGATGGCTGACACGAGCCCTGCGGGCATCAGCCCTCCAGCAAAATGGT 60
Db CTGAAGCTTCAAGATGGCTGACACGAGCCCTGCGGGCATCAGCCCTCCAGCAAAATGGT 108
QY 61 GGCCTCAGGCACCGGGCTGTGGTACCTCTCTCTCATGACACCCCTGGAGCTGGTGA 120
Db GGCCTCAGGCACCGGGCTGTGGTACCTCTCTCTCATGACACCCCTGGAGCTGGTGA 168
QY 121 GTTTCGCTGCAGTCTCAGCGCCCTCCATCGGCAGCGAGCTGATGCTTCTCCAGACT 180
Db GTTTCGCTGCAGTCTCAGCGCCCTCCATCGGCAGCGAGCTGATGCTTCTCCAGACT 228
QY 181 GTGAGGCTCTCTATACAAATAGCCTCTCTCTCCATCCACAGGAAAGTCCCTCT 240
Db GTGAGGCTCTCTATACAAATAGCCTCTCTCTCCATCCACAGGAAAGTCCCTCT 288
QY 241 GTATTGCAATGGTCTGAGGCTCTGTACCTGTGCCCCCAATGTGCGCTGTGCCAC 300
Db GTATTGCAATGGTCTGAGGCTCTGTACCTGTGCCCCCAATGTGCGCTGTGCCAC 348
QY 301 CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCGTGAG 360
Db CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCGTGAG 408
QY 361 GCACGAGGACACGAGACCTCTGAGGCGGCTCCCGGACCCCTGGTGTGATGATGTC 420
Db GCACGAGGACACGAGACCTCTGAGGCGGCTCCCGGACCCCTGGTGTGATGATGTC 468
QY 421 AGCTACCGCCCTCTACTGCTACCTGCTATGACCACTGAGGCTTCTCTGTGCTGCGAGC 480
Db AGCTACCGCCCTCTACTGCTACCTGCTATGACCACTGAGGCTTCTCTGTGCTGCGAGC 528
QY 481 CCTGACCTTGACCTCTACGCAACCATGGTGGTGGCGGCTGGCGCCGCTGGCCACCGT 540
Db CCTGACCTTGACCTCTACGCAACCATGGTGGTGGCGGCTGGCGCCGCTGGCCACCGT 588
QY 541 GACTGTGATCAGCCCTGGAGCTTATGCGGAACAGCTGCAAGCTTCAGAGTGTGTCGTA 600
Db GACTGTGATCAGCCCTGGAGCTTATGCGGAACAGCTGCAAGCTTCAGAGTGTGTCGTA 648
QY 601 CCGGAGCTGGTGGCTGTGTTGCACTGAGTGGCTCAGGCTGCTGCGCTCAGCTGCTG 660
Db CCGGAGCTGGTGGCTGTGTTGCACTGAGTGGCTCAGGCTGCTGCGCTCAGCTGCTG 708
QY 661 GCTGGCTGGGCCCCACTGCCCCCTCGAGATGTGCCCTTCTCAGCCCTGTACT-GGTCTCA 719
Db GCTGGCTGGGCCCCACTGCCCCCTCGAGATGTGCCCTTCTCAGCCCTGTACTGTCTCA 768
Q 720 ACTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGCTTCTGTGG 779
Db ACTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGCTTCTGTGG 828
QY 780 GCATGAGCTTGTGGCTGTGTCATCTCAGGAGCGTGTGCTGCTGCTGCTTACCT 839
Db GCATGAGCTTGTGGCTGTGTCATCTCAGGAGCGTGTGCTGCTGCTTACCT 888
QY 840 TTGAGCTGTGTAAGACCCCAAGCCAGGTGCTCT-GGGAGCGCATGGAGCTGTGAGAGTG 898
Db TTGAGCTGTGTAAGACCCCAAGCCAGGTGCTCT-GGGAGCGCATGGAGCTGTGAGAGTG 947
QY 899 AACCCCTGCATGTGAGCTCCACCTGGCTGCTGCTGCGAGATCGGGCCGAGTCGGG 958
Db AACCCCTGCATGTGAGCTCCACCTGGCTGCTGCTGCGAGATCGGGCCGAGTCGGG 1007
QY 959 ACCAAGGAGCTCT 971
Db ACCAAGGAGCTCT 1020
RESULT 3
AK007934
LOCUS
DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length

AK007934 1450 bp mRNA linear HTC 19-JAN-2002

enriched library, clone:1810062G21:homolog to HYPOTHETICAL 38.5 KDA
PROTEIN, full insert sequence.
AK007934 1 GI:12841807
VERSION
HTC; CAP trapper.
KEYWORDS
SOURCE
Mus musculus (strain:CS7Bl/6J) 10 day old male pancreas cDNA to
mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
clone:1810062G21.
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Resole, G.,
Quackenbush, J., Schriml, L. M., Scubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, X. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5 (bases 1 to 1450)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numata, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sasaki, K., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, I., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCGAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGAGCGCGCCCAATTAATTCGAGTTAATTAATTCGCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR

Location/Qualifiers
 1. 1450
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:1810062G21"
 /db_xref="MGD:MG1:1895562"
 /db_xref="taxon:10090"
 /clone="1810062G21"
 /sex="male"
 /tissue types="pancreas"
 /clone lib="RIKEN full-length enriched mouse cDNA library"
 /dev stage="10 day old"
 98. 1177
 /note="data source: SPTR, source key: Q9UF66, evidence: ISS homolog to HYPOTHETICAL 38.5 KDA PROTEIN putative"
 /codon start=1
 /protein id="BAR25360.1"
 /db_xref="GI:12841808"
 /db_xref="MGD:MG1:1915316"
 /translations="MDDQPGIGPLQOMVAGAGAVVTSLEMTPLDVVKVRLQSRP SATSELTSPSRFMSLSYTKSSALQSPKGLLYCNGVLEPLVCPNGTRCATWPDPT RFTGLDAFVKI VRHGGTITLWSGLPATVNTVPATAI VFTAYDOLKAFGLQSLTSD LYAPVAGALARMGTVTVVSPLVRLTKLQAHVSVRELASSVOAATVCGWSRLWLG WGPALRDVFPFSAFYWYVLYKSLRPLKQTSVTSVFSVAGISGMYAATLILFF DVVKTQRNSLGNVAERVRKPPRVDSWTLRLRRAESGTIRGLFAGFLPFIKAAPSC AIMISVEFKSFQRLNQEPQGR"

CDS
 /note="data source: SPTR, source key: Q9UF66, evidence: ISS homolog to HYPOTHETICAL 38.5 KDA PROTEIN putative"
 /dev stage="10 day old"
 98. 1177

BASE COUNT 267 a 456 c 402 g 325 t

Query Match 76.28; Score 848.6; DB 11; Length 1450;
 Best Local Similarity 85.98; Pred. No. 2.2e-195;
 Matches 941; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 2 TGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCATCAGCCCTCCAGCAAAATGGTG 61
 DB |||||
 QY 86 TGAATTCAGATGATGATCAGGACCTTGGGGCATTTAGCCCCCTTCAGCAAAATGGTG 145
 DB |||||
 QY 62 GCTCAGGACCGGGCTGTGGTACCTCTCTTCATGACACCCCTGGACGTGGTGAAG 121
 DB |||||
 QY 146 GCTCAGGACCGGGCTGTGGTACCTCTCTTCATGACACCCCTGGATGGTGAAG 205
 DB |||||
 QY 122 GTTCGCTCAGTCTCAGCGGCTTCATGCGCAGGAGCTGATGCTTCTCTCCAGACTG 181
 DB |||||
 QY 206 GTCCGCTTCACTCTCAGAGACCTCGGCAACAGCAATTCAGCACTCCCTCCAGATTC 265
 DB |||||

QY 182 TGGAGCCTCTCTATACCAAAATGGCCCTCTCTCTCAATCCACAGGAGATGGCTCTG 241
 DB |||||
 QY 266 TGGAGTCTCTCTTACACCAATATCTCTCCCTCTACAGTCCCGGGAAGTGGCTCTTA 325
 DB |||||
 QY 242 TATGCAATGGTGTCTTGAGGCTCTGTAACCTGTGCGCCAAATGGTGGCCGCTGTGCCACC 301
 DB |||||
 QY 326 TACTGCAATGGAGTCTCTGAGCGCCCTGTACCTGTGCGCAAAATGGTACCCTGTGCCACC 385
 DB |||||
 QY 302 TGGTTTCAAGACCTTACCGCTTCACTGGCACCATGATGCGCTTGTGAAGATGTGAGG 361
 DB |||||
 QY 386 TGGTTTCAAGACCTTACCGCTTCACTGGCACCATGATGCGCTTGTGAAGATGTGAGG 445
 DB |||||
 QY 362 CACAGGGCACAGGACCTCTGGAGGGGCTCTCCCGCCACCTCGTGTATGACTGTGCCA 421
 DB |||||
 QY 446 CATGAGGGCATTAGGACCTGTGGAGGGGCTCTCCAGCCACCTGTGTATGACCGTGCCA 505
 DB |||||
 QY 422 GCTACCGCATCTACTTCACTGCTGCTATGACCACTAAGGCTTCTCTGTGTGCTGAGCC 481
 DB |||||
 QY 506 GCTACTCTATCTACTTCACTGCTTACGACCAACTCAAGGCTTCTCTGTGTGCTCAGTCC 565
 DB |||||
 QY 482 CTGACCTCTGACCTTACGACCCATGTCGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 541
 DB |||||
 QY 566 TTGACCTCTGACCTTACGACCCATGTCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 625
 DB |||||
 QY 542 ACTGTGATCAGCCCTCTGAGGCTTATCGGACCAAGCTGACGCTCAGCATGTGTCTGATC 601
 DB |||||
 QY 626 ACATTTCTAGCCCTTGGAGCTGTGGGACCAAGCTGACGCTCAGCATGTGTCTATC 685
 DB |||||
 QY 602 OGGGAGCTGGTGGCTGTGTTGCACTGCACTGCTCAGGCTGGCTGGCTGGCTGGCTGG 661
 DB |||||
 QY 686 CGTGAGCTGGCTGTCTCTGTTCAAGCTGCGGTGACTCAGGCTGGCTGGCTGGCTGG 745
 DB |||||
 QY 662 CTGGCTGGGGCCCACTGCGCTTTCGAGATGTGCTCCCTTCTCAGCCCTCTACTGTTCAAC 721
 DB |||||
 QY 746 CTGGCTGGGGTCCACAGCTTCTCGAGATGTGCTCCCTTCTCAGCTCTGTACTGTTCAAC 805
 DB |||||
 QY 722 TATGAGCTGGTGAAGCTGGCTCAATGGGCTCAGCCGAGGACCAAGCTTCTGTGGGC 781
 DB |||||
 QY 806 TAGAGTTGGTGAAGAGCTGGCTGAGCGGACTGAGACCAAAAGACAGACATCGTGGGC 865
 DB |||||
 QY 782 ATGAGCTTGTGGCTGGTGGCATCTCAGGAGCGGTGGCTGCACTGCTGACTTACCTTT 841
 DB |||||
 QY 866 ATCAGCTTGTGGCGGTGGCATCTCAGGAAATGGTGGTGGCCACCTTACCTTACCTTTC 925
 DB |||||
 QY 842 GAGTGGTAAAGACCAACCGCAGGTGCTCTTGGAGCGATGGAGGCTGTGAGAGTGAAC 901
 DB |||||
 QY 926 GATGTGTTGAAGACACAGCGACAGATGTCACTGGGAGCGTGGAGGCTGTGAGAGTGAAG 985
 DB |||||
 QY 902 CCCCTGCATGTGGACTCCACCTGGCTGCTGCTGGAGGATCCGGGGCCGAGTCGGGACC 961
 DB |||||
 QY 986 CCGCCAGAGTGGATCCACCTGGCTCTGCTTCGAGAAATCCGGGCTGAAATCTGGCACC 1045
 DB |||||
 QY 962 AAGGACTCTTTTTCAGGCTTCTCTTCCTCGGATCATCAGGCTGCCCCCTCTCTGCGCATC 1021
 DB |||||
 QY 1046 AGGGACTCTTTTTCAGGCTTCTCTTCCTCCAGGATCATCAGGCTGACCTCTGTGGCATC 1105
 DB |||||
 QY 1022 ATGATCAGCACTTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAAACAGGACCGGCTT 1081
 DB |||||
 QY 1106 ATGATCAGCACTTATGATTCGGCAAAAGCTTCTTCCAGAGGCTCAACAGGAGAGCCT 1165
 DB |||||
 QY 1082 CTGGCGCTGGAAG 1096
 DB |||||
 QY 1166 CTGGCGCTGGAAG 1180
 DB |||||

RESULT 4
 AK007276 1371 bp mRNA linear HTC 19-JAN-2002
 LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700125C:8:homolog to HYPOTHETICAL 38.5 KDA PROTEIN, full insert sequence.
 DEFINITION AK007276
 ACCESSION AK007276
 VERSION AK007276.1 GI:12840721

KEYWORDS

SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:1700125C18.
 ORGANISM Mus musculus

REFERENCE

AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349836

REFERENCE

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE

AUTHORS Kawai,J., Shinegawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Konno,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojohji,T., Bono,H., Kasukawa,T., Saito,R., Kadoya,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Freischmann,M., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Tomita,M., Quackenbush,J., Schriml,L.M., Staupli,F., Suzuki,R., Tonita,H., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsi,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kaniwa,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weiss,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohrsuki,S. and Hayashizaki,Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851

REFERENCE

AUTHORS Adachi,J., Aizawa,K., Akahita,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohtno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinegawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 TITLE

TITLE

JOURNAL

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

FEATURES

Location/Qualifiers

1..1371
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:1700125C18"
 /db_xref="MGDI:1896541"
 /db_xref="taxon:10090"
 /clone="1700125C18"
 /sex="male"
 /tissue type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 123..1133
 /notes="data source:SEPT, source key:Q9UF66, evidence:ISS homolog to HYPOTHETICAL 38.5 KDA PROTEIN putative"

CDS

/codon_start=1
 /protein_id="BAB24932.1"
 /db_xref="GI:12840722"
 /db_xref="MGDI:1915316"
 /translation="WDQDQPMGTSPLQOMVAGAGAVTSLFMTPLDVVKVRLQSORP SATSELTSPRPMSLVSTKSSALOSPGLKLYCNGVLEPLVLCNPENCATWFOBPT RTGTDAFKIVHEGTRFLWGLPVTWVTPATATATFYATDQKAFGLCGSLTSD LVAPWVAGALARMGTIVVPSLELVRIKQAOHVSRELASSVQAVTQGGWKLWG WGTALRDPVSALYFNELVLSWLSGLSPKDOTSVGISFVAGGIGMVAATLTLPP DVVKTORQMSLGAVEAVEVRVPRVDSLTLLFAESGLNLAIPGDSLQASSPGSSRLHPV PS"
 BASE COUNT 255 a 437 c 367 g 292 t
 ORIGIN

Query Match 74.7%; Score 831.8; DB 11; Length 1371;
 Best Local Similarity 85.6%; Pred. No. 2.5e-191;
 Matches 937; Conservative 0; Mismatches 157; Indels 1; Gaps 1;

QY 2 TGAAGCTTCAAGATGGCTGACAGGACCTCGGGGATCATCGCCCTCCAGCAATGGTG 61
 DB 111 TGAATTCAGAGTATGATGATGAGACCTTTGGGGCATTAGCCCCCTTCAGCAATGGTG 170
 QY 62 GCCTCAGGACACCGGGCTGTGGTGTACCTCTCTTCATGACACCTCGGACCTGGTGAAG 121
 DB 171 GCCTCAGGACGCGGGCTGTGGTGTACCTCTCTTCATGACACCTCGGATGTGGTGAAG 230
 QY 122 GTTGGCTGAGTCTCAGCGGCTTCATGGCCAGGAGCTGATGCTTCTCCAGACTG 181
 DB 231 GTCCGCTTCAGTCTCAGAGACCTCGGCAACAGGCAATTGACAACCTCCCTCCAGATTG 290
 QY 182 TGGAGCTCTCTTATACCAATGGCTCTCTCTCTCAATCCACAGGAGTGCCTCTG 241
 DB 291 TGGAGTCTCTCTTACACCAATATCTCTCTCTCTCTCAGTCCCGGAGGAGTGCCTCTA 350
 QY 242 TATTGCAATGGTCTCTGGAGCTCTGTACCTGTGCGCCAAATGGTGGCGCTGTGCCACC 301
 DB 351 TACTGCAATGGAGTCTGGAGCCCTGTACTGTGCGCCAAATGGTACCGGTGTGCCACC 410
 QY 302 TGGTTTCAAGACCTTACCGCTTCACTGGCAATGATGCTTCTCGTAAGATCGTGAAG 361
 DB 411 TGGTTTCAAGACCTTACCGCTTCACTGGCAATGATGCTTCTCGTAAGATCGTGAAG 470
 QY 362 CACGAGGGCACCAGGACCTCTGGAGGGGCTCCCGCCACCTGGTGTGATGATGATGATG 421
 DB 471 CATGAGGGCCTAGGACCTCTGGAGGGGCTCCCGCCACCTGGTGTGATGATGATGATG 530
 QY 422 GCTACCGCCATCTTCTTCACTGCTCTATGACCAACTGAAGGCTTCTCTGTGTGCGAGCC 481

```

Db 531 GCTACTGCTATCTACTTCACTGCTTACGACCACTCAGGCTTCTGTGGTCACTCC 590
Qy 482 CTGACCTCTGACCTCTAGCCACCATGGTGGCTGGCCCTGGCCCTGGGACCGTG 541
Db 591 TTGACCTCTGACCTCTAGCCACCATGGTGGCTGGCCCTGGCCCTGGGACCGTG 650
Qy 542 ACTGTGATCAGCCCTGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTGTAC 601
Db 651 ACAGTTGTGAGCCCTTGGAGCTGTGGGACCAAGCTCAGGCTCAGCATGTGTGTAC 710
Qy 602 CGGAGCTGGGCTGCTGTGTTTGAAGTCTGAGTGGCTGAGGCTGGCTGAGTGG 661
Db 711 CGTGAAGCTGGCTGCTGTGTTTGAAGTCTGAGTGGCTGAGGCTGGCTGAGTGG 770
Qy 662 CTGGGCTGGGCTGGCTGCTGTGTTTGAAGTCTGAGTGGCTGAGGCTGGCTGAGTGG 721
Db 771 CTGGGCTGGGCTGGCTGCTGTGTTTGAAGTCTGAGTGGCTGAGGCTGGCTGAGTGG 830
Qy 722 TATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCTGAGGCTGAGGCTGAGGCT 781
Db 831 TAGAGTTGGTGAAGAGTGGCTCAATGGGCTCAGGCTGAGGCTGAGGCTGAGGCT 890
Qy 782 ATGAGCTGGTGGCTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 841
Db 891 ATGAGCTGGTGGCTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 950
Qy 842 GAGTGGTGAAGAGTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 901
Db 951 GATGGTGAAGAGTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 1010
Qy 902 CCCTGCTGATGGTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 961
Db 1011 CGGCTGATGGTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 1069
Qy 962 AAGGAGCTTGTGGAGCTGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 1021
Db 1070 AAGGAGCTTGTGGAGCTGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 1129
Qy 1022 ATGATCAGCCTATGAGTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 1081
Db 1130 ATGATCAGCCTATGAGTGGCTGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCT 1189
Qy 1082 CTGGGCTGGCTGAAAG 1096
Db 1190 CTGGGCTGGCTGAAAG 1204

```

```

RESULT: 5
AL522522
LOCUS AL522522 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF05 5
DEFINITION prime, mRNA sequence.
ACCESSION AL522522
VERSION AL522522.1 GI:12786015
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 891)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/gclone="CS0DB008YF05"
source
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

```

```

/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 146 a 292 c 264 g 189 t
ORIGIN

```

```

Query Match 74.1%; Score 826; DB 9; Length 891;
Best Local Similarity 99.9%; Pred. No. 5.2e-190;
Matches 837; Conservative 0; Mismatches 0; Indels 1; Gaps :
Qy 1 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCTATGAGCCCTCCAGCAATGCT 60
Db 55 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCTATGAGCCCTCCAGCAATGCT 114
Qy 61 GGCCTCAGGACACCGGGGCTGTGGTTACCTCTCTCTCATGACACCCCTCGACGTTGGA 120
Db 115 GGCCTCAGGACACCGGGGCTGTGGTTACCTCTCTCTCATGACACCCCTCGACGTTGGA 174
Qy 121 GGTTCGCTGAGTCTCAGGGGCTCTCCATGAGGAGCTGATGCTTCTCCAGACT 180
Db 175 GGTTCGCTGAGTCTCAGGGGCTCTCCATGAGGAGCTGATGCTTCTCCAGACT 234
Qy 181 GTGAGCCTCTCTATACCAATTTGCCCTCTCTCCCAATCGACAGGAGTGCCTCT 240
Db 235 GTGAGCCTCTCTATACCAATTTGCCCTCTCTCCCAATCGACAGGAGTGCCTCT 294
Qy 241 GTATTGCAATGGTGTCTGAGGCTCTGTACCTGTGCCCAATGGTGCCTGTGCCAC 300
Db 295 GTATTGCAATGGTGTCTGAGGCTCTGTACCTGTGCCCAATGGTGCCTGTGCCAC 354
Qy 301 CTGTTTCAAGACCTTACCGCTCTCATGACCATGATGCTTCTGTAAGATCGTAG 360
Db 355 CTGTTTCAAGACCTTACCGCTCTCATGACCATGATGCTTCTGTAAGATCGTAG 414
Qy 361 GCACGAGGACACGAGGACCTCTGAGAGCGGCTCTCCCGCACCTGATGATGCTGCC 420
Db 415 GCACGAGGACACGAGGACCTCTGAGAGCGGCTCTCCCGCACCTGATGATGCTGCC 474
Qy 421 AGCTACCGCATCTACTTCTGATGAGCAATGAGGCTTCTGTGGTGGAGC 480
Db 475 AGCTACCGCATCTACTTCTGATGAGCAATGAGGCTTCTGTGGTGGAGC 534
Qy 481 CTGACCTCTGACCTCTAGCACCCTATGGTGGCTGGGCGCTGGCCGCTGGGACCGT 540
Db 535 CTGACCTCTGACCTCTAGCACCCTATGGTGGCTGGGCGCTGGCCGCTGGGACCGT 594
Qy 541 GACTGTGATCAGCCCTCTGAGGCTTATGCGACAAAGCTGAGGCTGATGCTGCTGA 600
Db 595 GACTGTGATCAGCCCTCTGAGGCTTATGCGACAAAGCTGAGGCTGATGCTGCTGA 654
Qy 601 CCGGAGCTGGGCTGGCTGTGCTGAACTGAGTGGCTGAGGCTGGCTGCTGCTGCTG 660
Db 655 CCGGAGCTGGGCTGGCTGTGCTGAACTGAGTGGCTGAGGCTGGCTGCTGCTGCTG 714
Qy 661 GCTGGCTGGGCTGGCTGTGCTGAACTGAGTGGCTGAGGCTGGCTGCTGCTGCTG 720
Db 715 GCTGGCTGGGCTGGCTGTGCTGAACTGAGTGGCTGAGGCTGGCTGCTGCTGCTG 773
Qy 721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCTGAGGCTGAGGCTGCTG 780
Db 774 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCTGAGGCTGAGGCTGCTG 833

```

QY 781 CATGAGCTTTGGCTGTGTCATCTCAGGAGCGGTGGCTCAGTGTGACTTACCC 838
 Db 834 CATGAGCTTTGGCTGTGTCATCTCAGGAGCGGTGGCTCAGTGTGACTTACCC 891

RESULT 6
 BQ883556
 LOCUS
 DEFINITION AGENCOURT_8727656 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6340846
 5', mRNA sequence.
 ACCESSION BQ883556
 VERSION BQ883556.1 GI:22275564
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 900)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Tissue Procurement: NTCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LICM2537 row: 0 column: 23
 High quality sequence stop: 629.
 Location/Qualifiers
 1..900
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6340846"
 /clone_lib="NIH_MGC_47"
 /tissue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site: 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 150 a 286 c 262 g 200 t
 1..900

FEATURES
 source
 Query Match 74.1%; Score 826; DB 14; Length 900;
 Best Local Similarity 97.1%; Pred. No. 5.2e-190;
 Matches 873; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 116 GTGAGGTTTCGCTGAGTCTCAGCGGCTCCATGCGCAGGAGTGATGCTTCCTCC 175
 Db 1 GTGAGGTTTCGCTGAGTCTCAGCGGCTCCATGCGCAGGAGTGATGCTTCCTCC 60

QY 176 AGACTGTGGAGCTCTCCCTATACCAAAATGGCTCTCTCCATCCACAGGAAGTGC 235
 Db 61 AGACTGTGGAGCTCTCCCTATACCAAAATGGCTCTCTCCATCCACAGGAAGTGC 120

QY 236 CTCCTGTATTGCAATGGTGTCTGGAGGCTCTGTACCTGTGCCAAATGGTCCCGCTGT 295
 Db 121 CTCCTGTATTGCAATGGTGTCTGGAGGCTCTGTACCTGTGCCAAATGGTCCCGCTGT 180

QY 296 GCCACCTGGTTTCAAGACCTTACCGCTTCTACTGGACCATGATGCTTCGTGAAGATC 355
 Db 181 GCCACCTGGTTTCAAGACCTTACCGCTTCTACTGGACCATGATGCTTCGTGAAGATC 240

QY 356 GTGAGGCAACGAGGGCACCAGGACCTCTGGAGCGGCTCCCGCCACCCCTGGTGTGACT 415

Db 241 GTGAGGCAACGAGGGCACCAGGACCTCTGGAGCGGCTCCCGCCACCCCTGGTGTGACT 300
 QY 416 GTGCGAGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGGT 475
 Db 301 GTGCGAGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGGT 360
 QY 476 CGAGCCCTGACCTCTGACCTCTAGCGACCATGTTGGCTGGCGGCTGGCCCGCTGGGC 535
 Db 361 CGAGCCCTGACCTCTGACCTCTAGCGACCATGTTGGCTGGCGGCTGGCCCGCTGGGC 420
 QY 536 ACCGTGACTGTGATCAGCCCTCTGGAGCTTATCGCGACAAAGCTTCGAGGCTCAGCATGTG 595
 Db 421 ACCGTGACTGTGATCAGCCCTCTGGAGCTTATCGCGACAAAGCTTCGAGGCTCAGCATGTG 480
 QY 596 TCGTACCGGAGCTGGTGGCTGTGTGGAATGAGTGGCTGAGGCTGGCGCTCA 655
 Db 481 TCGTACCGGAGCTGGTGGCTGTGTGGAATGAGTGGCTGAGGCTGGCGCTCA 540
 QY 656 CTGTGCTGGGCTGGGCGCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGG 715
 Db 541 CTGTGCTGGGCTGGGCGCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGG 600
 QY 716 TTCACATATGAGCTGGTGAAGAGCTGGCTCAATGGCTCAGGCCGAAAGGACAGACTTCT 775
 Db 601 TTCACATATGAGCTGGTGAAGAGCTGGCTCAATGGCTCAGGCCGAAAGGACAGACTTCT 660
 QY 776 GTGGGATGAGCTTTGTGGCTGGTGGCATCTCAGGACGGTGGCTGAGTGTGACTCTA 835
 Db 661 GTGGGATGAGCTTTGTGGCTGGTGGCATCTCANGACGGTGGCTGAGTGTGACTCTA 720
 QY 836 CCTTTGAGCTGTGAAGACCCCAAGCCAGCTGCTCTGGAGGATGAGGCTGTGAGA 895
 Db 721 CCTTTGAGCTGTGAAGACCCCAAGCCAGCTGCTCTGGAGGATGAGGCTGTGAGA 780
 QY 896 GTGAACCCCTCGATGTGAGCTCCACCTGG-CTGCTGCTGGC-AGGATCCGGGCGAGT 953
 Db 781 GTGAACCCCTCGATGTGAGCTCCACCTGGCTGCTGCTGGAGAAACCCGGCAGT 840
 QY 954 CGGGACCAACAGGA---CTTTTGCAGGCTCTCTTCCTGCGATCATCAGGCTGCCCC 1009
 Db 841 CCGGCACCAAGGAACCTTTTTCAGGCTTCCTTTCTCGGATCATTTCAGGCTGCCCC 899

RESULT 7
 AL544360 917 bp mRNA linear EST 16-FEB-2001
 AL544360 LTI_NFL006 PL2 Homo sapiens cDNA clone CSODI018YN07 5
 prime, mRNA sequence.
 ACCESSION AL544360
 VERSION AL544360.1 GI:12876840
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 917)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1..917
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODI018YN07"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT)-primer. Five prime end

FEATURES
 source

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 153 a 293 c 268 g 195 t 8 others

ORIGIN

Query Match 73.4% Score 817.6; DB 9; Length 917;
Best Local Similarity 98.5%; Pred. No. 5.8e-188;
Matches 847; Conservative 8; Mismatches 2; Indels 3; Gaps 3;

QY 1 CTGAAGCTTCAGATGGCTGACAGGACCTCGGGGCAATGAGCCCTCCAGCAATGGT 60
DB 55 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCAATGAGCCCTCCAGCAATGGT 114
QY 61 GGCCTCAGGACCGGGGCTGTGTTACTCTCTCTTCATGACACCCCTGACGCTGTAA 120
DB 15 GGCCTCAGGACCGGGGCTGTGTTACTCTCTCTTCATGACACCCCTGACGCTGTAA 174
QY 121 GGTTCGCTGAGTCTGAGCGGCTCCATGCGCAGGAGCTGATGCTTCTCCAGACT 180
DB 175 GGTTCGCTGAGTCTGAGCGGCTCCATGCGCAGGAGCTGATGCTTCTCCAGACT 234
QY 181 GTGGAGGCTTCTATACCAATGGCTTCTCTCTCAATCCAGGAGGAGTCTCTCT 240
DB 235 GTGGAGGCTTCTATACCAATGGCTTCTCTCTCAATCCAGGAGGAGTCTCTCT 294
QY 241 GTATTGCAATGGTCTGAGGAGCTCTGACCTGTGCGCAATGGTGGCGCTGTCAC 300
DB 295 GTATTGCAATGGTCTGAGGAGCTCTGACCTGTGCGCAATGGTGGCGCTGTCAC 354
QY 301 CTGTTTCAAGACCTTACCGCTTCTGACGACCATGAGGCTTCTGAGGATGCTGAG 360
DB 355 CTGTTTCAAGACCTTACCGCTTCTGACGACCATGAGGCTTCTGAGGATGCTGAG 414
QY 361 GCACGAGGACGAGGAGGCTCTGAGGAGGCTCTCCGACCATGCTGATGATGCTGCC 420
DB 415 GCACGAGGACGAGGAGGCTCTGAGGAGGCTCTCCGACCATGCTGATGATGCTGCC 474
QY 421 AGCTACCGGCACTACTTCACTGCTATGACCAACTGAGGCTTCTGCTGCTGCTGAG 479
DB 475 AGACTACCGGCACTACTTCACTGCTATGACCAACTGAGGCTTCTGCTGCTGCTGAG 534
QY 480 CCCTGACCTTGACCTTACGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
DB 535 CCCTGACCTTGACCTTACGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
QY 540 TGACTGTGATCAGCCCTCGAGCTTATCGGACCAAGAGCTGAGGCTGAGCTGCTGT 599
DB 595 TGACTGTGATCAGCCCTCGAGCTTATCGGACCAAGAGCTGAGGCTGAGCTGCTGT 654
QY 600 ACCGAGGCTGGGTGCTGTGTTTCAACTGAGTGGCTCAGGCTGCTGCTGCTGCTGCT 659
DB 655 ACCGAGGCTGGGTGCTGTGTTTCAACTGAGTGGCTCAGGCTGCTGCTGCTGCTGCT 714
QY 660 GCGTGGGCTGGGCGCCACTGCGCTTCAGATGTCAGCTTCTCAGCCCT-EPACTGGTTC 718
DB 715 GCGTGGGCTGGGCGCCACTGCGCTTCAGATGTCAGCTTCTCAGCCCTGGTGGTGGTTC 774
QY 719 AACTATGAGCTGTGAGGAGCTGGCTCAATGGCTCAGGCTGAGGACGAGCTTCTGTG 778
DB 775 AACTATGAGCTGTGAGGAGCTGGCTCAATGGCTCAGGCTGAGGACGAGCTTCTGTG 833
QY 779 GGCATGAGCTTGTGGCTGGTGGCTTCTCAGGAGGCTGCTGCTGCTGCTGCTGCTGCT 838
DB 834 GGCATGAGCTTGTGGCTGGTGGCTTCTCAGGAGGCTGCTGCTGCTGCTGCTGCTGCT 893
QY 839 TTTGAGCTGTGTAAGACCA 858
|||||

Db 894 TTTGACCTGTGTAAGACCA 913

RESULT 8

AK002381

LOCUS

DEFINITION

AK002381 1453 bp mRNA linear HTC 19-JAN-2002
Mus musculus adult male kidney cDNA, RIKEN full-length enriched
library, clone:0610009E11:homolog to HYPOTHETICAL 38.5 KDA PROTEIN,
full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK002381.1 GI:12832321
HTC; CAP trapper.
Mus musculus (strain: C57BL/6J) adult male kidney cDNA to mRNA,
clone:lib:RIKEN full-length enriched mouse cDNA library
clone:0610009E11.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunuma, T., Tsunashima, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20499374
11042159

3

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, F., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nika, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staudli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schomburg, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 885-890 (2001)

21085660

11217851

5 (bases 1 to 1453)

AUTHORS

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,

Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yananaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGGAGAGAGCGCGCGCACTCGAGTGTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to R₀ = 5.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGGAGAGAGAGATCCAGAGCTCAATTAATTAACCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES source

Location/Qualifiers
1..1453
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:0610009E11"
/db_xref="MGI:MGI:1892017"
/db_xref="taxon:10090"
/clone="0610009E11"
/sex="male"
/tissue type="kidney"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1453
/note="data source:SPTR, source key:Q9UF66, evidence:ISS homolog to HYPOTHETICAL 38.5 KDA PROTEIN"
/db_xref="MGI:MGI:1915316"
268 a 456 c 403 g 326 t

Query Match 72.9%; Score 812.6; DB 11; Length 1453;
Best Local Similarity 85.7%; Pred. No. 1.2e-186;
Matches 938; Conservative 0; Mismatches 154; Indels 3; Gaps 3;

```

Qy 2 TGAAGCTTCAAGATGGCTGACAGAGCCCTGGGGATCAGCCCCCTCCAGCAATGGTG 61
Db 90 TGAATTCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 149

Qy 62 GCCTCAGGCACCGGGCTGTGGTACTCTCTCTTCATGACACCCCTGGACGTGGTGAAG 121
Db 150 GCCTCAGGAGCGGGCTGTGGTACTCTCTCTTCATGACACCCCTGGATGGTGAAG 209

Qy 122 GTTCGCTCAGTCTCAGCGCCCTCCATGGCCAGCAGCTGATGCTCTCTCCACACTG 181
Db 210 GTCCGCTTCTCAGTCTCAGAGACCTCGGCAACAGCGAATGACAACTCCCTCCAGATC 269

Qy 182 TGAAGCTTCTCTATACCAATGTCCTCTCTCTCCAAATCCAGGGAAGTCCCTCTG 241
Db 270 TGGAGTCTCTCTACCAATATCTCTCGCTCTACAGTCCCGAGGGAAGTGCCTCTTA 329

```

```

Qy 242 TATTCAATGGTGTCTGGAGCCCTGTACCTGTGCCCCAAATGGTGGCCGTGTGCACC 301
Db 330 TACTGAATGAGTCTGGAGCCCTGTACCTGTGCCCCAAATGGTGGCCGTGTGCACC 389
Qy 302 TGGTTTCAAGACCTTACCCGCTTCACTGGCACCATTGATGCTTCTGTGAAGATCTGAGG 361
Db 390 TGGTTTCAAGACCTTACCCGCTTCACTGGCACCATTGATGCTTCTGTGAAGATCTGCGG 449
Qy 362 CACGAGGACACGAGACCTTGGAGCGCCCTCCCGCCACCTCCCTGTGATGATGTGCA 421
Db 450 CATGAGGACATAGGACCTTGGAGCGCCCTCCCGCCACCTCCCTGTGATGATGTGCA 509
Qy 422 GCTACCGCATCTACTTCACTGCTTATGACCAATGAGGCTTCTGTGTGTGCGAC 481
Db 510 GCTACTGTACTTCACTGCTTACGACCAATCAAGGCTTCTGTGTGTGCGAC 569
Qy 482 CTGACCTTCACTGCTTACGACCAATGAGGCTTGGCGCTTGGCGCCCTGGGCACTG 541
Db 570 TTGACCTTCACTGCTTACGACCAATGAGGCTTGGCGCTTGGCGCCCTGGGCACTG 628
Qy 542 ACTGTGATCAGCCCTTGGAGCTTATGCGGCAAAAGCTGAGGCTCAGCATGTGCTGAC 601
Db 629 ACAGTGTGAGCCCTTGGAGCTTGGCGCAAAAGCTGAGGCTCAGCATGTGCTGAC 688
Qy 602 CCGAGCTGGGTGCTGTGTTGAACTGCTGCTCAGGCTGCTGCGGCTGGGCTCACTG 661
Db 689 CGTGAAGTGGGTGCTGTGTTGAACTGCTGCTCAGGCTGCTGCGGCTGGGCTCTCTG 748
Qy 662 CTGGGCTGGGGCCCTCACTGCTTCAAGATGCTGCTTCTCAGCCCTGTGCTGCTCAAC 721
Db 749 CTGGGCTGGGGCCCTCACTGCTTCAAGATGCTGCTTCTCAGCTCTGTGCTGCTCAAC 807
Qy 722 TATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCTGAGGCAAGGACAGACTTCTG 781
Db 808 TACAGTGTGAGAGCTGGCTCAGGCTGAGGCAAGGACAGACTTCTGAGGCTGGG 867
Qy 782 ATGAGCTTTTGGCTGGGCTGCTCAGGAGCGGTGCTGCTGCTGCTGCTGCTGCT 841
Db 868 ATGAGCTTTTGGCTGGGCTGCTCAGGAGCGGTGCTGCTGCTGCTGCTGCTGCT 927
Qy 842 GACGTGTAAGACCAACCGCAGGTGCTTCTGGAGCGGATGAGGCTGTGAGAGTGAAC 901
Db 928 GATGTGTGAAGACCAACCGCAGGTGCTTCTGGAGCGGATGAGGCTGTGAGAGTGAAG 987
Qy 902 CCCCTGATGCTGAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
Db 988 CCGCCAGAGTGAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
Qy 962 AAGGGAATCTTTGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
Db 1047 AAGGGAATCTTTGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Qy 1022 ATGATCAGCACTTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGCT 1081
Db 1107 ATGATCAGCACTTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGCT 1166
Qy 1082 CTGGCGGCTGAAAG 1096
Db 1167 CTGGCGGCTGAAAG 1181

```

RESULT 9
B0683596
LOCUS
DEFINITION
B0683596
AGENCY 8059897 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091198
5', mRNA sequence.
B0683596
VERSION
B0683596.1 GI:21796275
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
AUTHORS      1 (bases 1 to 920)
TITLE        NIH-MGC http://mgi.nci.nih.gov/
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/STP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2333 row: m column: 23
High quality sequence stop: 697.
Location/Qualifiers
1..920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6091198"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GCCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley), using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT  161 a 288 c 273 g 196 t      2 others
ORIGIN

Query March 72.2%; Score 803.8; DB 14; Length 920;
Best Local Similarity 96.8%; Pred. NO. 1.3e-184;
Matches 850; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGGT 50
DB 31 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGGT 90
QY 61 GGCTCAGGACCGGGCTGTGGTACTCTCTTTCATGACACCCCTGGACGTGGTAA 120
DB 91 GGCTCAGGACCGGGCTGTGGTACTCTCTTTCATGACACCCCTGGACGTGGTAA 150
QY 121 GGTTCGCTCGAGCTCTAGCGGCGCTCCATGGCCAGGCTGATGCTTCTCCAGACT 180
DB 151 GGTTCGCTCGAGCTCTAGCGGCGCTCCATGGCCAGGCTGATGCTTCTCCAGACT 210
QY 181 GTGGAGCTCTCTATACCAATTTGCCCTCTCTCTCCAAATCCAGGGAAGTGCCTCT 240
DB 211 GTGGAGCTCTCTATACCAAT-----GGAGTGCCTCT 246
QY 241 GTATTCGAATGGTCTCGAGGCTCTGTACTGTGCCAAATGGTCCCGCTGTGCCAC 300
DB 247 GTATTCGAATGGTCTCGAGGCTCTGTACTGTGCCAAATGGTCCCGCTGTGCCAC 306
QY 301 CTGGTTTCAAGACCTACCCGCTTCACTGSCACCATGATGCTTGTGAGATGCTGAG 360
DB 307 CTGGTTTCAAGACCTTACCCGCTTCACTGSCACCATGATGCTTGTGAGATGCTGAG 366
QY 361 GCACGAGGCGACCGAGCCCTCTGGAGCGGCTCCCGCCACCTGCTGATGACTGTGCC 420
DB 367 GCACGAGGCGACCGAGCCCTCTGGAGCGGCTCCCGCCACCTGCTGATGACTGTGCC 426
QY 421 AGCTACCGGCTACTTCTGCTGCTATGACCACTGAAGGCTTCTGTGTGCTGAGC 480
DB 427 AGCTACCGGCTACTTCTGCTGCTATGACCACTGAAGGCTTCTGTGTGCTGAGC 486
QY 481 CCTGACCTCTGACCTCTACGCCACCATGGTGGCGCTGGCCGCTGGGACCGT 540
DB 487 CCTGACCTCTGACCTCTACGCCACCATGGTGGCGCTGGCCGCTGGGACCGT 546

```

```

QY 541 GACTGTGATCAGCCCTGGAGCTTATCGGACAAAGCTGAGGCTCAGATGTGCTCA 600
DB 547 GACTGTGATCAGCCCTGGAGCTTATCGGACAAAGCTGAGGCTCAGATGTGCTCA 606
QY 601 CCGGAGCTGGGTGCTGTGTTCGAATCGAGTGGCTCAGGCTGGCTGGCGCTCACTGTG 660
DB 607 CCGGAGCTGGGTGCTGTGTTCGAATCGAGTGGCTCAGGCTGGCTGGCGCTCACTGTG 666
QY 661 GCTGGGCTGGGCCCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTTACTGTTCAA 720
DB 667 GCTGGGCTGGGCCCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTTACTGTTCAA 726
QY 721 CTATGAGCTGTGAAGAGTGGCTCAATGGGCTCAGGCGGAGGACCACTTCTGTGGG 780
DB 727 CTATGAGCTGTGAAGAGTGGCTCAATGGGCTCAGGCGGAGGACCACTTCTGTGGG 786
QY 781 CATGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGCGAGTCTTACCTACCC-T 839
DB 787 CATGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGCGAGTCTTACCTACCC-T 846
QY 840 TTGACGTGTGTAAGAGACCCCAACCCGAGCTGCTCTGGGA 877
DB 847 TTGACGTGTGTAAGAGACCCCAACCCGAGCTGCTCTGGGA 884
RESULT 10
LOCUS      AK019396      1491 bp      mRNA      linear      HTC 19-JAN-2002
DEFINITION Mus musculus 12 days embryo head cDNA, RIKEN full-length enriched
            library, clone:3010027G13:homolog to HYPOTHETICAL 38.5 KDA PROTEIN,
            full insert sequence.
ACCESSION  AK019396
VERSION    AK019396.1 GI:12859581
KEYWORDS   HTC; CAP trapper.
SOURCE     Mus musculus (strain:CE7BL/6J) 12 days embryo head cDNA to mRNA,
            cDNA library
            Cloned from: RIKEN full-length enriched mouse cDNA library
            Cloned from: 3010027G13.
ORGANISM   Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1
AUTHORS    Carninci, P. and Hayashizaki, Y.
TITLE      High-efficiency full-length cDNA cloning
JOURNAL    Meth. Enzymol. 303, 19-44 (1999)
MEDLINE    99279253
PUBMED     10349636
REFERENCE  2
AUTHORS    Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE      Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL    Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE    20499374
PUBMED     11042159
REFERENCE  3
AUTHORS    Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
            Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwara, S., Inoue, K., Iizawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE      RIKEN integrated sequence analysis (RISA) system-384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL    Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE    20530913
PUBMED     11076861
REFERENCE  4
AUTHORS    Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
            Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamahata, I.,
            Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

```


Search completed: January 23, 2003, 00:15:26
Job time : 1772.67 secs